

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 9.63934 Seconds

(without alignments)
279,413 Million cell updates/sec

Title: US-09-673-166B-273
Perfect score: 140
Sequence: 1 GRQYKANSKFITGTERGSLNVTATL 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	50.0	1315	1 BTCLTN	lentoxilysin (EC 3
2	52	37.1	500	2 S33979	gag polypotein -
3	52	37.1	836	2 T18650	hypothetical prote
4	50	35.7	478	1 FOYWL	gag polypotein -
5	50	35.7	497	1 FOLJND	gag polypotein -
6	50	35.7	498	2 T09436	gag polypotein -
7	50	35.7	500	1 FOYWL	gag polypotein -
8	50	35.7	500	1 A44001	gag polypotein -
9	50	35.7	500	1 FOYWL	gag polypotein -
10	50	35.7	502	1 FOYWL	gag polypotein -
11	50	35.7	512	1 FOYWL	gag polypotein -
12	49	35.0	506	1 A38068	gag polypotein -
13	49	35.0	506	1 A55236	gag polypotein -
14	48.5	34.6	1060	2 S06286	gag polypotein -
15	48.5	34.6	1086	2 S16752	gag polypotein -
16	48.5	34.6	1701	2 A54498	gag polypotein -
17	48.5	34.6	1701	2 A26868	gag polypotein -
18	48.5	34.6	1726	2 SAZOGM	gag polypotein -
19	48.5	34.6	1726	2 A45948	gag polypotein -
20	48.5	34.6	3216	2 C90538	gag polypotein -
21	48	34.3	83	2 B84174	gag polypotein -
22	48	34.3	355	2 A5318	gag polypotein -
23	48	34.3	975	2 A5318	gag polypotein -
24	48	34.3	1126	2 T05761	gag polypotein -
25	47.5	33.9	1126	2 T05761	gag polypotein -
26	47	33.6	505	2 T01667	gag polypotein -
27	47	33.6	569	2 B82559	gag polypotein -
28	47	33.6	674	2 T20571	gag polypotein -
29	47	33.6	998	2 I41078	hemolysin - Escher

30	47	33.6	998	2 T00227	hemolysin A toxin
31	47	33.6	1161	2 G81186	conserved hypotet
32	47	33.6	1161	2 G81186	hypothetical prote
33	46	32.9	97	2 H64335	hypothetical prote
34	46	32.9	131	2 T12915	hypothetical prote
35	46	32.9	191	2 C64001	hypothetical prote
36	46	32.9	349	2 T43043	probable acetyl-Co
37	46	32.9	402	2 D81061	conserved hypotet
38	46	32.9	404	2 B81071	conserved hypotet
39	46	32.9	422	2 F81804	hypothetical prote
40	46	32.9	424	2 G97834	hemolysin homolog
41	46	32.9	447	2 H97146	hemolysin/Surfac
42	46	32.9	448	2 B64503	hypothetical prote
43	46	32.9	501	2 S54377	gag polypotein -
44	46	32.9	823	2 G89007	protein P59B1.8 [I
45	46	32.9	999	2 S68689	glucose regulated

ALIGNMENTS

RESULT 1
BTCLTN
lentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N/Alternate names: tetanus neurotoxin
C/Species: Clostridium tetani
C/Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C/Accession: A25689; A25757; A25194; A60759; S69348; S09364
R/Eisel, U.; Jarasch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M
EMBL J. 5, 2495-2502, 1986
A>Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botu
A/Reference number: A25689; MUID:87053814; PMID:3536478
A/Accession: A25689
A/Molecule type: DNA
A/Residues: 1-1315 <EIS>
A/Cross-references: GB:X04436; NID:G40769; PIDN:CAA28033.1; PID:G40770
A/Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A>Title: The complete nucleotide sequence of tetanus toxin.
A/Reference number: A25757; MUID:87040747; PMID:3774547
A/Accession: A25757
A/Molecule type: DNA
A/Residues: 1-1315 <FAI>
A/Cross-references: GB:X06214; NID:G40773; PIDN:CAA2564.1; PID:G40774
A/Experimental source: strain CN3911
R/Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A>Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in E
A/Reference number: A25194; MUID:86085672; PMID:3510187
A/Accession: A25194
A/Molecule type: DNA
A/Residues: 743-1315 <FA2>
A/Cross-references: GB:M12739; NID:G144920; PIDN:AAA23282.1; PID:G144921
A/Accession: B25194
A/Molecule type: Protein
A/Residues: 865-894 <PA3>
R/Matsuda, M.; Ieli, D.L.; Sugimoto, N.; Ozutsami, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A>Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal
A/Reference number: A60759; MUID:90035436; PMID:2478476
A/Accession: A60759
A/Molecule type: Protein
A/Residues: 461-475 <MAT>
J. Demotz, S.; Iannavechia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 144, 394-402, 1989
A>Title: Determination of several DR-restricted tetanus toxin T cell epitopes.
A/Reference number: J50098; MUID:89093918; PMID:2463305
A/Contents: annotation; epitope region
R/Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.R
Nature 359, 832-835, 1992
A>Title: Tetanus and botulinum B neurotoxins block neurotransmitter release by proteolytic
A/Reference number: S27125; MUID:93063293; PMID:1331807
A/Contents: annotation

R.de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A>Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A/Reference number: S69348; MUID:95262688; PMID:7744050
A/Accession: S69348
A/Molecule type: protein
A/Residues: 2-31 <DEP>
C/Comment: The source of this protein was an extrachromosomal plasmid.
C/Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C/Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglionic
C/Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C/Function:
A/Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C/Superfamily: tetanus toxin
C/Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F/2-457/Product: tetroxylisin light chain (fragment A) #status predicted <TTH>
F/461-1315/Product: tetroxylisin heavy chain (fragment B,C) #status experimental <TTH>
F/461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F/865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F/233-237/Binding site: zinc (His) #status predicted
F/234/Active site: Gln #status predicted

Query Match 50.0%; Score 70; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QYKANSKFIGITE 16
DB 830 QYKANSKFIGITE 843

RESULT 2
333979
gag polyprotein - human immunodeficiency virus type 1
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C/Accession: S33979
R/Carlini, F.
Submitted to the EMBL Data Library, November 1991
A/Reference number: S33979
A/Accession: S33979
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-500 <CAR>
A/Cross-references: EMBL:Z11530; NID:G60192; PIDN:CAA77621.1; PID:G60193
C/Superfamily: AIDS-related virus gag polyprotein

Query Match 37.1%; Score 52; DB 2; Length 500;
Best Local Similarity 78.6%; Pred. No. 7.8;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 72 SEERSLYNTVATL 85

RESULT 3
T18650
Hypothetical protein B0035.12 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T18650
R/White, S.
Submitted to the EMBL Data Library, May 1996
A/Reference number: Z19002
A/Accession: T18650
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-836 <WIL>
A/Cross-references: EMBL:Z73102; PIDN:CAA97405.1; GSPDB:GN00022; CESP:B0035.12
A/Experimental source: clone B0035

C/Genetics:
A/Gene: CESP:B0035.12
A/Map position: 4
A/Introns: 28/3; 78/2; 134/3; 167/1; 299/3; 349/2; 550/1; 662/3; 700/3; 799/1

Query Match 37.1%; Score 52; DB 2; Length 836;
Best Local Similarity 52.2%; Pred. No. 13;
Matches 12; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 KANSKFIGTERGRSLYNTVATL 28
DB 342 RARSNVINSAEGRSLYNTVATL 364

RESULT 4
FOVWVL
gag polyprotein - human immunodeficiency virus type 1 (isolate LV)
N/Alternate names: assemblin; core polyprotein; gag precursor
N/Contains: capsid antigen core protein p24CA; core protein p1; core protein p6; matrix ;
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 05-Dec-1998
C/Accession: A03948
R/Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A/Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi
A/Reference number: A93355; MUID:8511157; PMID:2982104
A/Accession: A03948
A/Molecule type: DNA
A/Residues: 1-478 <MOE>
C/Genetics:
A/Gene: gag
C/Superfamily: AIDS-related virus gag polyprotein
C/Keywords: AIDS; blocked amino end; core protein; immunodeficiency; lipoprotein; myrist
F/2-478/Product: gag precursor (assemblin) #status predicted <GAG>
F/2-132/Product: matrix antigen core protein p17MA #status predicted <P17>
F/20-32/Region: nuclear location signal
F/110-114/Region: nuclear location signal
F/133-163/Product: capsid antigen core protein p24CA #status predicted <P24>
F/366-377/Product: core protein p2 #status predicted <CP2>
F/378-452/Product: nucleocapsid core protein p7NC #status predicted <CP7>
F/392-405/Region: zinc finger CCHC motif
F/413-426/Region: zinc finger CCHC motif
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F/392,395,400,405/Binding site: zinc (Cys, Cys, His, Cys) #status experimental
F/413,416,421,426/Binding site: zinc (Cys, Cys, His, Cys) #status experimental

Query Match 35.7%; Score 50; DB 1; Length 478;
Best Local Similarity 78.6%; Pred. No. 15;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
DB 72 SEERSLYNTVATL 85

RESULT 5
POLJND
gag polyprotein - human immunodeficiency virus type 1 (isolate NDK)
N/Alternate names: core polyprotein
N/Contains: core protein p15; core protein p17; core protein p24
C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C/Accession: J00065
R/Spire, B.; Site, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
Gene 81, 275-284, 1989
A/Title: Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human immunc
A/Reference number: J00065; MUID:90034200; PMID:2806917
A/Accession: J00065
A/Molecule type: DNA
A/Residues: 1-497 <SPI>
A/Cross-references: GB:M27323; NID:G328154; PIDN:AAA44868.1; PID:G328157
C/Genetics:

```
433-448/Product: core protein p1 #status predicted <CP1>
```

Produced by NETOS

Produced by NETOS

Best Local Similarity 78.6%; Pred. No. 16;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
: |||||
Db 72 SEELRSLYNTVATL 85

RESULT 10

gag polyprotein - human immunodeficiency virus type 1 (isolate ARV-2)

N/Alternate names: core polyprotein

N/Contains: core protein p15; core protein p17; core protein p24

C/Species: human immunodeficiency virus type 1, HIV-1

A/Note: host Homo sapiens (man)

C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C/Accession: A03947

R/Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-SH

Science 227, 484-492, 1985

A/Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).

A/Reference number: A04003; MUID:85090453; PMID:2578227

C/Comment: Cleavage sites that yield the mature core proteins remain to be determined.

C/Genetics:

A/Genes: gag

C/Superfamily: AIDS-related virus gag polyprotein

C/Keywords: AIDS; core protein; immunodeficiency; polyprotein

F/1-134/Product: core protein p17 #status predicted <P17>

F/135-193/Product: core protein p24 #status predicted <P24>

F/394-502/Product: core protein p15 #status predicted <P15>

Query Match 35.7%; Score 50; DB 1; Length 502;
Best Local Similarity 78.6%; Pred. No. 16;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
: |||||
Db 72 SEELRSLYNTVATL 85

RESULT 11

gag polyprotein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

N/Alternate names: core polyprotein

N/Contains: core protein p15; core protein p17; core protein p24

C/Species: human immunodeficiency virus type 1, HIV-1

A/Note: host Homo sapiens (man)

C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C/Accession: A03945

R/Ratner, L.; Haseltine, W.; Patarca, R.; Liyak, K.J.; Starcich, B.; Josephs, S.F.; Dora

nderger, J.A.; Papas, T.S.; Ghayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

Nature 313, 277-284, 1985

A/Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A/Reference number: A9353; MUID:85111123; PMID:2578615

A/Accession: A03945

A/Molecule type: DNA

A/Residues: 1-512 <RAT>

A/Cross-references: GB:M1654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA442

C/Comment: Cleavage sites that yield the mature core proteins remain to be determined.

C/Genetics:

A/Genes: gag

C/Superfamily: AIDS-related virus gag polyprotein

C/Keywords: AIDS; core protein; immunodeficiency; polyprotein

F/1-132/Product: core protein p17 #status predicted <P17>

F/133-391/Product: core protein p24 #status predicted <P24>

F/392-512/Product: core protein p15 #status predicted <P15>

QY 15 TERGRSLYNTVATL 28
: |||||
Db 72 SEELRSLYNTVATL 85

RESULT 12

gag polyprotein - human immunodeficiency virus type 1 (strain MN)

N/Alternate names: core polyprotein

N/Contains: core protein p1; core protein p17; core protein p2; core protein p24; core p1

C/Species: human immunodeficiency virus type 1, HIV-1

A/Note: host Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Feb-1999

C/Accession: A38068

R/Henderson, L.B.; Bowers, M.A.; Sowder II, R.C.; Serabyn, S.A.; Johnson, D.G.; Bees Jr.,

J. Virol. 66, 1856-1865, 1992

A/Title: Gag proteins of the highly replicative MN strain of human immunodeficiency virus

A/Reference number: A38068; MUID:92194415; PMID:1548743

A/Accession: A38068

A/Molecule type: protein

A/Residues: 1-506 <HEN>

C/Genetics:

A/Genes: gag

C/Superfamily: AIDS-related virus gag polyprotein

C/Keywords: AIDS; core protein; immunodeficiency; polyprotein

F/1-134/Product: core protein p17 #status experimental <P17>

F/135-193/Product: core protein p24 #status experimental <P24>

F/366-379/Product: core protein p2 #status experimental <P2>

F/380-434/Product: core protein p7 #status experimental <P7>

F/435-450/Product: core protein p1 #status experimental <P1>

F/451-506/Product: core protein p6 #status experimental <P6>

Query Match 35.0%; Score 49; DB 1; Length 506;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
: |||||
Db 71 SEELRSLYNTVATL 84

RESULT 13

A55236

kinesin-related protein KLP68D - fruit fly (Drosophila melanogaster)

N/Alternate names: kinesin-like protein 5; KLP5

C/Species: Drosophila melanogaster

C/Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 19-Jan-2001

C/Accession: A55236; E41298

R/Pesavento, P.A.; Stewart, R.J.; Goldstein, L.S.B.

J. Cell Biol. 127, 1041-1048, 1994

A/Title: Characterization of the KLP68D kinesin-like protein in Drosophila: possible role

A/Reference number: A55236; MUID:95050960; PMID:7525600

A/Accession: A55236

A/Molecule type: mRNA

A/Residues: 1-784 <PEB>

A/Cross-references: GB:U15974; NID:9595912; PIDN:AAA69929.1; PID:9565090

R/Stewart, R.J.; Pesavento, P.A.; Woerpel, D.N.; Goldstein, L.S.B.

Proc. Natl. Acad. Sci. U.S.A. 88, 8470-8474, 1991

A/Title: Identification and partial characterization of six members of the kinesin super

A/Reference number: A41298; MUID:92020874; PMID:1924306

A/Accession: E41298

A/Molecule type: DNA

A/Residues: 'TC', '222-337', 'VRGOV' <STE>

A/Cross-references: GB:M74431; NID:9157791; PIDN:AAA28658.1; PID:9157792

C/Genetics:

A/Genes: FlyBase:KLP68D; KLP5

A/Cross-references: FlyBase:FBgn0004381

C/Function:

A/Description: may be part of a motor protein that provides anterograde fast axonal tran

C/Superfamily: kinesin-related protein KIF3; kinesin motor domain homology

C/Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop

F/1-349/Domain: head globular #status predicted <GGL>

F:20-350/Domain: kinesin motor domain homology <KMO>
 F:106-113/Region: nucleotide-binding motif A (P-loop)
 F:350-580/Domain: helical rod #status predicted <ROD>
 F:581-784/Domain: tail globular #status predicted <TEL>
 F:112/Binding site: ATP (lys) #status predicted

Query Match 35.0%; Score 49; DB 1; Length 784;
 Best Local Similarity 42.9%; Pred. No. 36;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 2 RQYIKANSKFEIGTERGRSLY 22
 Db 163 RDLKPKSKHLEVRERGGVY 183

RESULT 14
 S06286
 major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (st
 N;Alternate names: 190K protein; polymorphic schizont antigen
 C;Species: Plasmodium falciparum
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jun-2000
 C;Accession: S06286
 R;Certa, U.; Rotmann, D.; Matile, H.; Reber-Liske, R.
 EMBL J 6 4137-4142, 1987
 A;Title: A naturally occurring gene encoding the major surface antigen precursor p190 of
 A;Reference number: S06286; MUID:98166657; PMID:3327688
 A;Accession: S06286
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-1060 <CER>
 C;Superfamily: major merozoite surface antigen
 C;Keywords: surface antigen

Query Match 34.6%; Score 48.5; DB 2; Length 1060;
 Best Local Similarity 43.5%; Pred. No. 58;
 Matches 10; Conservative 7; Mismatches 5; Indels 1; Gaps 1;
 QY 2 RQYIKANSKFEIGTERGRSLY 23
 Db 982 KQFVKNSKVITGLTGTQKNALN 1004

RESULT 15
 S16752
 major merozoite surface antigen MSA-1 - malaria parasite (Plasmodium falciparum) (fragme
 N;Alternate names: polymorphic schizont antigen p190
 C;Species: Plasmodium falciparum
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jun-2000
 C;Accession: A44865; S16752
 R;Olafsson, P.; MacIle, H.; Certa, U.
 Exp. Parasitol. 74, 381-389, 1992
 A;Title: Plasmodium falciparum: the repetitive MSA-1 surface protein of the RO-71 isolat
 A;Reference number: A44865; MUID:92275047; PMID:1592091
 A;Accession: A44865
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1086 <OLA>
 A;Cross-references: EMBL:X61930
 A;Experimental source: isolate RO-71
 C;Genetics:
 A;Gene: MSA1
 C;Superfamily: major merozoite surface antigen
 C;Keywords: glycoprotein; merozoite; surface antigen; tandem repeat

Query Match 34.6%; Score 48.5; DB 2; Length 1086;
 Best Local Similarity 43.5%; Pred. No. 59;
 Matches 10; Conservative 7; Mismatches 5; Indels 1; Gaps 1;
 QY 2 RQYIKANSKFEIGTERGRSLY 23
 Db 1008 KQFVKNSKVITGLTGTQKNALN 1030

Search completed: July 20, 2004, 06:34:08
 Job time : 11.6393 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 5.5082 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166b-273
Perfect score: 140
Sequence: 1 GRQYKXNSKFIETGERSLVNTVATL 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	70	50.0	1314	1	TETX_CLOTE
2	54	38.6	499	1	GAG_HV1E1
3	51	36.4	499	1	GAG_HV1L1
4	51	36.4	513	1	MAD2_RHIL1
5	51	36.4	562	1	MM25_HOMAN
6	50	35.7	388	1	GAG_HV1W2
7	50	35.7	496	1	GAG_HV1OY
8	50	35.7	499	1	GAG_HV1C4
9	50	35.7	499	1	GAG_HV1H2
10	50	35.7	499	1	GAG_HV1H2
11	50	35.7	511	1	GAG_HV1A2
12	50	35.7	511	1	GAG_HV1B1
13	50	35.7	511	1	GAG_HV1B5
14	50	35.7	511	1	GAG_HV1B1
15	50	35.7	511	1	GAG_HV1B1
16	50	35.7	511	1	GAG_HV1B1
17	49	35.0	492	1	GAG_HV1U4
18	49	35.0	506	1	GAG_HV1U4
19	49	35.0	784	1	KL68_DROME
20	48.5	34.6	1682	1	MSP1_PLA3
21	48.5	34.6	1701	1	MSP1_PLA3
22	48.5	34.6	1701	1	MSP1_PLA3
23	48.5	34.6	1726	1	MSP1_PLA3
24	48.5	34.6	1726	1	MSP1_PLA3
25	47	33.6	504	1	GAG_HV1A1
26	47	33.6	532	1	TP6B_SULTO
27	46	32.9	97	1	TP6B_SULTO
28	46	32.9	191	1	Y096_HA2IN
29	46	32.9	448	1	Y096_HA2IN
30	46	32.9	500	1	Y096_HA2IN
31	46	32.9	999	1	Y096_HA2IN
32	46	32.9	999	1	Y096_HA2IN
33	45	32.1	111	1	RU24_CHLNU

34	45	32.1	117	1	YC01_YEAST
35	45	32.1	146	1	DUT_TREPA
36	45	32.1	151	1	Y228_AQUAE
37	45	32.1	152	1	Y228_AQUAE
38	45	32.1	152	1	Y228_AQUAE
39	45	32.1	152	1	Y228_AQUAE
40	45	32.1	152	1	Y228_AQUAE
41	45	32.1	152	1	Y228_AQUAE
42	45	32.1	152	1	Y228_AQUAE
43	45	32.1	152	1	Y228_AQUAE
44	45	32.1	152	1	Y228_AQUAE
45	45	32.1	152	1	Y228_AQUAE

ALIGNMENTS

RESULT 1	ID	TETX_CLOTE	STANDARD	PRT	1314 AA.
AC	P04958				
DT	13-AUG-1987	(Rel. 05, Created)			
DT	13-AUG-1987	(Rel. 05, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) (Contains: Tetanus toxin light chain [Tetanus toxin chain L]; Tetanus toxin heavy chain [Tetanus toxin chain H]).				
DE	Tetanus toxin heavy chain [Tetanus toxin chain H].				
GN	TETX OR CTPE0.				
OG	Clostridium tetani.				
OG	Plasmid pE88, and Plasmid 75 Kbp.				
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;				
OC	Clostridium.				
OX	NCBI_TaxID=1513;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	PLASMID=75 Kbp;				
RC	MEDLINE=87053814; PubMed=3536478;				
RA	Eisel U., Jarasch W., Goretzki K., Henschen A., Engels J.,				
RA	Weller U., Hadel M., Habermann B., Niemann H.;				
RT	"Tetanus toxin: Primary structure, expression in E. coli, and				
RT	homology with botulinum toxins.";				
RT	EMBO J. 5:2495-2502(1986).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CN911; PLASMID=75 Kbp;				
RC	MEDLINE=87040747; PubMed=3774547;				
RA	Faltweather N.F., Lyness V.A.;				
RA	"The complete nucleotide sequence of tetanus toxin.";				
RT	Nucleic Acids Res. 14:7809-7812(1986).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Massachusetts / E88; PLASMID=PE88;				
RC	MEDLINE=22457253; PubMed=12552129;				
RA	Brieglebmann H., Baumer S., Fricke W.F., Wietzer A., Liesegang H.,				
RA	Decker I., Herzberg C., Martinez-Arias R., Merl R., Henne A.,				
RA	Gottschalk G.;				
RT	"The genome sequence of Clostridium tetani, the causative agent of				
RT	tetanus disease.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).				
RL	[4]				
RP	SEQUENCE OF 742-1314 FROM N.A.				
RC	PLASMID=75 Kbp;				
RC	MEDLINE=86085672; PubMed=3510187;				
RA	Faltweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;				
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin				
RT	fragment C in Escherichia coli.";				
RT	J. Bacteriol. 165:21-27(1986).				
RL	[5]				
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.				
RC	MEDLINE=90201034; PubMed=2108021;				
RA	Kriegstein K., Henschen A., Weller U., Habermann E.;				
RT	"Arrangement of disulfide bridges and positions of sulfhydryl groups				
RT	in tetanus toxin.";				

RL Eur. J. Biochem. 188:39-45(1990).
RN [6]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites."
RL Eur. J. Biochem. 202:41-51(1991).
RN [7]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc."
RL EMBJ. 11:3577-3583(1992).
RN [8]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=131807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin."
RL Nature 359:832-835(1992).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin."
RL Nat. Struct. Biol. 4:788-792(1997).
CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-PHE-77
CC BOND OF SYNAPTOSOMAL-2.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-Phe-77 bond in
CC synaptobrevin 2.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GLYCOSYLTRANSFERASES.
CC -1- SIMILARITY: Belongs to peptidase family M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X0436; CAA28033.1; -;
CC EMBL: X06214; CAA29564.1; -;
CC EMBL: AF528097; AA037454.1; -;
CC EMBL: M12739; AAA23282.1; -;
CC PIR: A25689; BTCLTN.
CC PDB: 1AF9; 29-APR-98.
CC PDB: 1ABD; 14-OCT-98.
CC PDB: 1DOH; 27-MAR-00.
CC PDB: 1DPO; 24-MAR-00.
CC PDB: 1DIW; 24-MAR-00.
CC PDB: 1DLI; 24-MAR-00.
CC PDB: 1FV3; 05-SEP-01.
CC MEROPS: M27.001; -;
CC InterPro: IPR008985; Cons like Jec_g1.
CC InterPro: IPR002160; Kunitz_legume.
CC InterPro: IPR006025; Pept_M_Zn_BS.

DR InterPro: IPR00395; Peptidase_M27.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolyase; Metalloprotease; Zinc; Plasmid;
KW 3D-structure; Complete proteome.
FT INIT MET 0
FT CHAIN 1
FT METAL 457
FT ACT SITE 232
FT METAL 233
FT METAL 236
FT TRANSMEM 226
FT TRANSMEM 669
FT DISULFID 438
FT DISULFID 1076
FT HELIX 876
FT TURN 883
FT STRAND 884
FT TURN 892
FT STRAND 894
FT STRAND 904
FT TURN 909
FT STRAND 912
FT STRAND 920
FT TURN 928
FT STRAND 932
FT HELIX 938
FT TURN 941
FT STRAND 949
FT HELIX 962
FT TURN 972
FT STRAND 980
FT STRAND 983
FT STRAND 987
FT TURN 996
FT STRAND 998
FT TURN 1006
FT STRAND 1010
FT STRAND 1020
FT TURN 1021
FT STRAND 1031
FT TURN 1039
FT STRAND 1042
FT TURN 1047
FT STRAND 1048
FT STRAND 1050
FT TURN 1058
FT STRAND 1068
FT TURN 1079
FT STRAND 1082
FT HELIX 1097
FT TURN 1106
FT STRAND 1112
FT STRAND 1114
FT TURN 1116
FT STRAND 1120
FT STRAND 1122
FT TURN 1123
FT STRAND 1127
FT STRAND 1131
FT HELIX 1132
FT TURN 1135
FT STRAND 1141
FT TURN 1144
FT STRAND 1148
FT STRAND 1152
FT STRAND 1155
FT TURN 1159
FT STRAND 1163
FT STRAND 1173
FT TURN 1184
FT STRAND 1188
FT STRAND 1190
TETANUS TOXIN LIGHT CHAIN.
ZINC (CATALYTIC) (BY SIMILARITY).
BY SIMILARITY.
ZINC (CATALYTIC) (BY SIMILARITY).
POTENTIAL.
INTERCHAIN.

Query Match 50.0%; Score 70; DB 1; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QYKANSKFIGITE 16
 DB 829 QYKANSKFIGITE 842

RESULT 2
 GAG_HV1EL
 ID GAG_HV1EL STANDARD; PRT; 499 AA.
 AC P04592;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polypeptide [Contains: Core proteins p17, p24, p2, p7, p1, p6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (ELI isolate) (HTV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
 OX NCBI_TaxID=11689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86245056; PubMed=2424612;
 RA Allison M., Main-Hobson S., Montagnier L., Sonigo P.;
 RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
 of two isolates from African patients";
 RL Cell 46:63-74 (1986).
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -1- PTM: The p24 protein is phosphorylated.
 CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K03454; AAA44324.1; -
 DR EMBL; A07108; CA000611.1; -
 DR HSP; P05888; 1AAR.
 DR HIV; K03454; GAGSHEL.
 DR InterPro; IPR000721; Gag_p24.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.
 DR InterPro; IPR000071; Retrovir_p17.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_p24; 1.
 DR Pfam; PF00098; Zf_CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00243; HIVMATRIX.
 DR SMART; SM00343; ZNF_C2HC; 2.
 DR SMART; SM00343; ZNF_C2HC; 2.
 DR PROSITE; PS50158; ZF_CCHC; 2.
 DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
 Zinc-finger; Repeat; Lipoprotein.
 KM Zinc-finger; Repeat; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 131
 FT CHAIN 132 362
 FT CHAIN 363 377
 FT CHAIN 378 431
 FT CHAIN 432 447
 FT CHAIN 448 499
 FT CHAIN 499 499
 FT ZN_FING 390 407
 FT CCHC-TYPE 1.

FT ZN_FING 411 428 CCHC-TYPE 2.
 FT LIPID 1 1 N-myristoyl glycine (in host) (By
 FT SEQUENCE 499 AA; 55582 MW; 8A1785A59EARD08D CRC64;
 SO

Query Match 38.6%; Score 54; DB 1; Length 499;
 Best Local Similarity 85.7%; Pred. No. 1.4;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 DB 71 TERGRSLYNTVATL 84

RESULT 3
 GAG_HV1LW
 ID GAG_HV1LW STANDARD; PRT; 499 AA.
 AC Q70622;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polypeptide [Contains: Core proteins p17, p24, p2, p7, p1, p6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HTV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentiviruses.
 OX NCBI_TaxID=82834;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95127297; PubMed=7826699;
 RA Reitz M.S., Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
 RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
 RT "Viral variability and serum antibody response in a laboratory worker
 infected with HIV type 1 (HTLV type IIIB).";
 RL AIDS Res. Hum. Retroviruses 10:1143-1155 (1994).
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -1- PTM: The p24 protein is phosphorylated.
 CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U12055; AAA76686.1; -
 DR PDB; 1HVN; 31-JAN-94.
 DR PDB; 1HVO; 31-JAN-94.
 DR PDB; 2ZNF; 15-JUL-92.
 DR InterPro; IPR000721; Gag_p24.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.
 DR InterPro; IPR000071; Retrovir_p17.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_p24; 1.
 DR Pfam; PF00098; Zf_CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00243; HIVMATRIX.
 DR SMART; SM00343; ZNF_C2HC; 2.
 DR PROSITE; PS50158; ZF_CCHC; 2.
 DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
 Zinc-finger; Repeat; 3D-structure; Lipoprotein.
 KM Zinc-finger; Repeat; 3D-structure; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 131
 FT CHAIN 132 362
 FT CHAIN 362 362
 FT CORE PROTEIN P17 (MATRIX PROTEIN).
 FT CORE PROTEIN P24 (CORE ANTIGEN).

```

FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7. (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P1.
FT CHAIN 448 499 CORE PROTEIN P6.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 1 N-myristoyl glycine (in host) (By
FT SEQUENCE 499 AA; 55772 MW; AA2F683546DC0A9 CRC64;
  similarity).
Query Match 36.4%; Score 51; DB 1; Length 499;
Best Local Similarity 78.6%; Pred. No. 4.1;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGSRSLYNTVATL 28
Db 71 SEGRSLYNTVATL 84

RESULT 4
NADB_RHILO STANDARD; PRT; 513 AA.
AC 0984TV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE L-aspartate oxidase (EC 1.4.3.16) (LASPO) (Quinolinate synthetase B).
NADB OR ML5834.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
[1]
SEQUENCE FROM N.A.
RP STRAIN=MAFF303099;
RC MEDLINE=21082930; PubMed=11214968;
RA Kato T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RT DNA Res. 7:331-338(2000).
[2]
FUNCTION: Catalyzes the oxidation of L-aspartate to
iminoaspartate.
-1- CATALYTIC ACTIVITY: L-aspartate + H(2)O + O(2) = oxaloacetate +
NH(3) + H(2)O(2).
-1- COFACTOR: FAD.
-1- PATHWAY: NAD biosynthesis; aspartate to NAMN; first step.
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-1- SIMILARITY: BELONGS TO THE FAD-DEPENDENT OXIDOREDUCTASE FAMILY 2.
NADB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
-----
CC EMBL, AP003007; BAB52214.1; -
DR InterPro; IPR003953; FAD_bind2.
DR InterPro; IPR005288; NADB.
DR InterPro; IPR004112; Succ_DH flav_C.
DR Pfam; PF00890; FAD_binding_2; 1.
DR Pfam; PF02910; succ_DH flav_C; 1.
DR TIGRPFAM; TIGR00551; nadb; 1.
KW Pyridine nucleotide biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW Complete proteome.
FT NP_BIND 13 27 FAD (AMP PART) (POTENTIAL).
FT ACT_SITE 233 233 BY SIMILARITY.

```

```

FT ACT SITE 252 252 BY SIMILARITY.
SQ SEQUENCE 513 AA; 53469 MW; 6DA2B4B63F4200BC CRC64;
Query Match 36.4%; Score 51; DB 1; Length 513;
Best Local Similarity 55.0%; Pred. No. 4.2;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 SKRIGTERGRSLYNTVATL 28
Db 422 SNALGIVRNKSLCDIVATL 441

RESULT 5
NM25 HUMAN STANDARD; PRT; 562 AA.
ID NM25_HUMAN;
AC Q9NP42; Q9H3Q0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Matrix metalloproteinase-25 precursor (EC 3.4.24.-) (MMP-25)
DE (Membrane-type matrix metalloproteinase 6) (MT6-MMP 6) (Membrane-type-6
DE matrix metalloproteinase) (MT6-MMP) (Leukolysin).
GN MMP25 OR MT6MMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Fetal liver;
RC MEDLINE=2016815; PubMed=10706098;
RA Velasco G., Cal S., Merlos-Suarez A., Ferrando A.A., Alvarez S.,
RA Nakano A., Arribas J., Lopez-Otin C.;
RA "Human MT6-matrix metalloproteinase: identification, progelatinase A
RT activation, and expression in brain tumors.";
RT Cancer Res. 60:877-882(2000).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=20092460; PubMed=10628838;
RA Pei D.O.;
RA "Leukolysin/MMP25/MT6-MMP: a novel matrix metalloproteinase
RT specifically expressed in the leukocyte lineage.";
RT Cell Res. 9:291-303(1999).
[3]
SEQUENCE FROM N.A.
RP Kojima S., Itoh Y., Matsumoto S., Seiki M.;
RA "Membrane-type 6 matrix metalloproteinase (MT6-MMP, MMP-25) is the
RT second glycosyl-phosphatidyl inositol (GPI)-anchored MMP.";
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RP De Santis-Vis B.M., Clair-Moniot V.A., Lambert C.A., Vandervliet B.,
RA Pin J.O., Chalus L., Alt-Yahia S., Caux C., Richelle-Nusgens B.V.,
RA Fossiez F., Lebecque S.;
RA "Molecular cloning of a novel human membrane-type matrix
RT metalloproteinase (MT-MMP) predominantly expressed in dendritic
RT cells.";
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[5]
GPI-ANCHOR.
RP MEDLINE=20487129; PubMed=11034316;
RA Kojima S.-I., Itoh Y., Matsumoto S.-I., Masuno Y., Seiki M.;
RA "Membrane-type 6 matrix metalloproteinase (MT6-MMP, MMP-25) is the
RT second glycosyl-phosphatidyl inositol (GPI)-anchored MMP.";
RA FEBS Lett. 480:142-146(2000).
[6]
FUNCTION: May activate progelatinase A.
-1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI anchor.
-1- TISSUE SPECIFICITY: Expressed predominantly in leukocytes, lung
CC and spleen. Expressed also in colon carcinoma, astrocytoma and
CC glioblastomas.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).

```

```

CC -1- SIMILARITY: Belongs to peptidase family M10A.
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ239053; CAB94713.1; -
DR EMBL; AF145442; AAP6697.2; -
DR EMBL; AF185270; AAG17007.1; -
DR EMBL; AB042328; BAB20584.1; -
DR EMBL; AJ272137; CAC03490.1; -
DR HSSP; Q02853; 1HV5.
DR MEROPS; M10.024; -.
DR Genew; HGNC:14246; MMP25.
DR GO; GO:0005578; C:extracellular matrix; NAS.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0006994; P:inflammatory response; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; peptidase_M10; 1.
DR PRINTS; PR00138; Peptidase_M10_N; 1.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZINC; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydroxylase; Metalloprotease; Zinc; Calcium; Signal; Zymogen;
KW GFI-anchor; Extracellular matrix; Lipoprotein.
FT SIGNAL 1 21
FT PROPEP 22 107 BY SIMILARITY.
FT CHAIN 108 539 MATRIX METALLOPROTEINASE-25.
FT PROPEP 540 562 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 314 511 HEMOPEXIN-LIKE.
FT SITE 90 90 CYSTEINE SWITCH (POTENTIAL).
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 234 234 BY SIMILARITY.
FT METAL 237 237 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 243 243 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 317 508 BY SIMILARITY.
FT LIPID 539 539 GPI-anchor amidated alanine (Potential).
FT DOMAIN 103 107 POLY-ARG.
FT CHAIN 549 555 POLY-LEU.
FT DOMAIN 47 47
FT CONFLICT 562 AA; 62554 MM; A6A50AE05D969C64;
SQ SEQUENCE 562 AA; 62554 MM; A6A50AE05D969C64;
Query Match 36.4%; Score 51; DB 1; Length 562;
Best Local Similarity 44.4%; Pred. No. 4.7;
Matches 12; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
QY 2 RQYIYANSKFGITRGSLVNTVATL 28
Db 59 RDAIKVMORFAGLPETGRMDPGTVATM 85

```

```

OS Human immunodeficiency virus type 1 (MMJ2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCHI_taxid=11705;
OX
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal P., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT at risk for AIDS."
RL Science 232:1548-1553(1986).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC -1- BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- P1M: The p24 protein is phosphorylated.
CC -1- MISCELLANEOUS: ISOLATES MMJ1, MMJ2, AND MMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03457; AAB12988.1; -
DR PDB; 1FGJ; 01-APR-97.
DR HIV; K03457; GAG$MMJ2.
DR InterPro; IPR000721; Gag_p24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR PRINTS; PR00234; HIVMATRIX.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; 3D-structure; Lipoprotein.
FT INIT MET 0 0 BY SIMILARITY.
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 >388 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT LIPID 1 1 N-myristoyl glycine (in host) (by
FT similarity).
FT NON TER 388 388
SQ SEQUENCE 388 AA; 43314 MM; EF885BE10ECF7804 CRC64;
Query Match 35.7%; Score 50; DB 1; Length 388;
Best Local Similarity 78.6%; Pred. No. 4.5;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGSLVNTVATL 28
Db 71 SEELSLVNTVATL 84

```

```

RESULT 6
GAG_HYI1W2
ID GAG_HYI1W2 STANDARD; PRT; 388 AA.
AC P05889;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6]
DE (Fragment).
GN GAG.

```

```

RESULT 7
GAG_HYI1ND
ID GAG_HYI1ND STANDARD; PRT; 496 AA.
AC P18600;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

NCBI	Taxid=11695;
[1]	
SEQUENCE FROM N.A.	
MEDLINE=90034200; PubMed=2806917;	
Spire B., Spire U., Zachar V., Rey F., Barre-Sinoussi F., Galibert F., Hampe A., Chermann J.C.;	
"Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human immunodeficiency virus.";	
RT Gene 81:275-284(1989).	
- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED STAGES DURING THE ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.	
GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.	
- P1M: The p24 protein is phosphorylated.	
- MISCELLANEOUS: NDK ISOLATED FROM A ZAIREAN PATIENT AFFECTED WITH AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.	
- SIMILARITY: Contains 2 CCHC-type zinc fingers.	
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outbreak - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).	
EMBL; M27323; AAAA4868.1; -. PIR; J00065; FOLDND. PDB; 1EQI; 30-SEP-94. HIV; M27323; GAGSNDK. InterPro; IPR00721; Gag_p24. InterPro; IPR008916; Retrov_capsid_C. InterPro; IPR008919; Retrov_capsid_N. InterPro; IPR000701; Retrovir_p17. InterPro; IPR001878; Znf_CCHC. Pfam; PF00540; Gag_p17; 1. Pfam; PF00607; Gag_p24; 1. Pfam; PF00098; ZF_CCHC; 2. PRINTS; PR00939; C2HCZNFINGER. PRINTS; PR00234; HIVMATRIX. SMART; SMO0343; Znf_C2HC; 2. PROSITE; PS50158; ZF_CCHC; 2. AIDS; Core protein; Polyprotein; Myristate; Phosphorylation; Zinc-finger; Repeat; 3D-structure; Lipoprotein. BY SIMILARITY.	
INIT MET 0 128 CORE PROTEIN P17 (MATRIX PROTEIN).	
CHAIN 129 359 CORE PROTEIN P24 (CORE ANTIGEN).	
CHAIN 360 374 CORE PROTEIN P2.	
CHAIN 375 449 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).	
CHAIN 430 445 CORE PROTEIN P1.	
CHAIN 446 496 CORE PROTEIN P6.	
ZN FING 387 404 CCHC-TYPE 1.	
ZN FING 408 425 CCHC-TYPE 2.	
LIPID 1 1 N-myristoyl glycine (in host) (By similarity).	
SEQUENCE 496 AA; 55152 MW; F50M42E42CCFE20AA CRC64;	
Query Match 35.7%; Score 50; DB 1; Length 496;	
Best Local Similarity 78.6%; Pred. No. 5.8;	
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0	
OY 15 TERGRSLVNTVATL 28	
:	
DB 71 SEIRSLVNTVATL 84	
RESULT 8	
GAG_HV10Y	
ID _GAG_HV10Y STANDARD; PRT; 498 AA.	
P20889;	

QY	DB	15	TERGRLSYNTVATL 28	71	SEIRLSYNTVATL 84	Score 50; DB 1; Length 498; PRed. No. 5.9; Indels 0; Gaps 0;
QY	DB	15	TERGRLSYNTVATL 28	71	SEIRLSYNTVATL 84	Score 50; DB 1; Length 498; PRed. No. 5.9; Indels 0; Gaps 0;


```

RESULT 9
GAG_HV1C4
ID GAG_HV1C4 STANDARD; PRT; 499 AA.
AC P05887;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polypeptide [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11687;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87041461; PubMed=3490666;
RA Desai S.M., Kalyanaraman V.S., Casey J.M., Srinivasan A.,
RA Andersen P.R., Devare S.G.;
RT "Molecular cloning and primary nucleotide sequence analysis of a
RT distinct human immunodeficiency virus isolate reveal significant
RT divergence in its genomic sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: The p24 protein is phosphorylated.
CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13136; AAA44306.1; -.
DR PIR; A25523; FOVMH4.
DR HSSP; P05888; 1A4F.
DR HIV; M13136; GAGSCDC45.
DR InterPro; IPR0000721; Gag_p24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR Pfam; PF00607; Zf_CCHC; 2.
DR PRINTS; PR00098; Zf_CCHC; 2.
DR PRINTS; PR00939; C2HC2NFINGER.
DR SMART; SM00343; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PS50158; Zf_CCHC; 2.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; Lipoprotein.
FT INIT MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2.
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 448 499 CORE PROTEIN P1.
FT ZN_FING 389 406 CORE PROTEIN P6.
FT ZN_FING 410 427 CCHC-TYPE 1.
FT LIPID 1 1 N-myristoyl glycine (in host) (by
SQ SEQUENCE 499 AA; 55796 MW; 023CA76C9C6F22AD CRC64;
Query Match 35.7%; Score 50; DB 1; Length 499;
Best Local Similarity 78.6%; Pred. No. 5.9;

```

```

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 15 TERGRSLYNTVATL 28
Db 71 SEELRSLYNTVATL 84

RESULT 10
GAG_HV1H2
ID GAG_HV1H2 STANDARD; PRT; 499 AA.
AC P04591;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polypeptide [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87299196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -1- PTM: The p24 protein is phosphorylated.
CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50258.1; -.
DR PDB; 1B6T; 11-DEC-00.
DR PDB; 1ESK; 21-APR-00.
DR PDB; 1NCP; 31-OCT-93.
DR PDB; 1TAM; 12-NOV-96.
DR HIV; K03455; GAGSHXB2.
DR InterPro; IPR0000721; Gag_p24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovir_p17.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR PRINTS; PR00098; Zf_CCHC; 2.
DR PRINTS; PR00939; C2HC2NFINGER.
DR SMART; SM00343; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PS50158; Zf_CCHC; 2.
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; 3D-structure; Lipoprotein.
FT INIT MET 0
FT CHAIN 1 131 BY SIMILARITY.
FT CHAIN 132 362 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 363 376 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 377 431 CORE PROTEIN P2.
FT CHAIN 432 447 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 448 499 CORE PROTEIN P1.
FT ZN_FING 389 406 CORE PROTEIN P6.
FT ZN_FING 410 427 CCHC-TYPE 1.

```

FT ZN_FING 410 427 CCHC-TYPE 2.
 FT LIPID 1 1 N-myristoyl glycine (in host) (By
 FT SEQUENCE 499 AA; 55798 MW; 774C384D6EACB108 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 499;
 Best Local Similarity 78.6%; Pred. No. 5.9;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGSLYNTVATL 28
 Db 71 SEELRSLYNTVATL 84

RESULT 11
 GAG_HV1A2 STANDARD; PRT; 499 AA.
 AC P35962;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93021387; PubMed=1404605;
 RA L.Y. Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
 Shaw G.M.;
 RT "Complete nucleotide sequence, genome organization, and biological
 RT properties of human immunodeficiency virus type 1 in vivo: evidence
 RT for limited defectiveness and complementation.";
 RL J. Virol. 66:6587-6600 (1992).
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -1- P17: The p24 protein is phosphorylated.
 CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M93258; -; NOT_ANNOTATED_CDS.
 CC PIR: A44001; A44001.
 CC PDB: 1FEU; 09-OCT-00.
 CC InterPro: IPR000721; Gag_p24.
 CC InterPro: IPR008916; Retrov_capsid_C.
 CC InterPro: IPR008919; Retrov_capsid_N.
 CC InterPro: IPR000071; Retrov_p17.
 CC InterPro: IPR001878; Znf_CCHC.
 CC Pfam: PF00540; Gag_p17; 1.
 CC Pfam: PF00607; Gag_p24; 1.
 CC PRINTS: PR000939; C2HCZNFINGER.
 CC PRINTS: PR00234; HIVMATRIX.
 CC SMART: SM00343; ZNF_C2HC; 2.
 CC PROSITE: PS50158; ZF_CCHC; 2.
 CC ALD5; Core protein; Polyprotein; Myristate; Phosphorylation;
 CC zinc-finger; Repeat; 3D-structure; Lipoprotein.
 CC INIT MET 0
 CC CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
 FT

FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
 FT CHAIN 363 376 CORE PROTEIN P2 (NUCLEOCAPSID PROTEIN).
 FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
 FT CHAIN 432 447 CORE PROTEIN P1.
 FT CHAIN 448 499 CORE PROTEIN P6.
 FT ZN_FING 389 406 CCHC-TYPE 1.
 FT ZN_FING 410 427 CCHC-TYPE 2.
 FT LIPID 1 1 N-myristoyl glycine (in host) (By
 FT SEQUENCE 499 AA; 55660 MW; 278B65F5405CD99 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 499;
 Best Local Similarity 78.6%; Pred. No. 5.9;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGSLYNTVATL 28
 Db 71 SEELRSLYNTVATL 84

RESULT 12
 GAG_HV1A2 STANDARD; PRT; 501 AA.
 AC P03349;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85090453; PubMed=2578227;
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steiner K.S.,
 Stempelen M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
 Levy J.A., Dina D., Luciw P.A.;
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
 RT (ARV-2).";
 RL Science 227:484-492 (1985).
 CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -1- P17: The p24 protein is phosphorylated.
 CC -1- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: K02007; AAB59875.1; -.
 CC PIR: A03947; FOVMA2.
 CC HSSP: P05888; 1AAF.
 CC HIV: K02007; GAGSF2.
 CC InterPro: IPR000721; Gag_p24.
 CC InterPro: IPR008916; Retrov_capsid_C.
 CC InterPro: IPR008919; Retrov_capsid_N.
 CC InterPro: IPR000071; Retrov_p17.
 CC InterPro: IPR001878; Znf_CCHC.
 CC Pfam: PF00540; Gag_p17; 1.
 CC Pfam: PF00607; Gag_p24; 1.
 CC Pfam: PF00098; ZF_CCHC; 2.
 CC PRINTS: PR00939; C2HCZNFINGER.
 CC PRINTS: PR00234; HIVMATRIX.
 DR

DR SMART: SM00343; ZNF C2HC; 2.
 DR PROSITE; PS50158; ZF CCHC; 2.
 KW AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
 FT ZINC-FINGER; Repeat; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 133 CORE PROTEIN P17 (MATRIX PROTEIN).
 FT CHAIN 134 364 CORE PROTEIN P24 (CORE ANTIGEN).
 FT CHAIN 365 378 CORE PROTEIN P2.
 FT CHAIN 379 433 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
 FT CHAIN 434 449 CORE PROTEIN P1.
 FT CHAIN 450 501 CORE PROTEIN P6.
 FT ZN FING 391 408 CCHC-TYPE 1.
 FT ZN FING 412 429 CCHC-TYPE 2.
 FT LIPID 1 N-myristoyl glycine (in host) (By similarity).
 SQ SEQUENCE 501 AA; 55935 MW; 8F36C928D5EDBA45 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 501;
 Best Local Similarity 78.6%; Pred. No. 5.9;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 15 TERGRSLYNTVATL 28
 Db 71 SEELRSLYNTVATL 84
 ID GAG_HV1B1 STANDARD; PRT; 511 AA.
 AC P03347;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
 RA Lautenberger J.A., Pappas T.S., Ghayab J., Chang N.T., Gallo R.C.,
 RA Wong-Staal F.;
 "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 Nature 313:277-284(1985).
 RT Nature 313:277-284(1985).
 RL Nature 313:277-284(1985).
 CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -!- PTM: The p24 protein is phosphorylated.
 CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M15654; AAA44201.1; -
 DR PIR; A03945; F0VWH3.
 DR HSP; P05888; IAAF.
 DR HIV; M15654; GAGSBH102.
 DR InterPro; IPR000721; Gag_p24.
 DR InterPro; IPR008916; Retrov_capsid_C.
 DR InterPro; IPR008919; Retrov_capsid_N.

DR InterPro; IPR000721; Retrovit p17.
 DR InterPro; IPR001878; Znf CCHC.
 DR Pfam; PF00540; Gag_p17; 1.
 DR Pfam; PF00607; Gag_p24; 1.
 DR Pfam; PF00098; zf-CCHC; 2.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR PRINTS; PR00234; HIVMATRIX.
 DR SMART; SM00343; ZNF C2HC; 2.
 DR PROSITE; PS50158; ZF CCHC; 2.
 KW AIDS; Core protein; Polypeptide; Myristate; Phosphorylation;
 FT ZINC-FINGER; Repeat; Lipoprotein.
 FT INIT MET 0
 FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
 FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
 FT CHAIN 363 376 CORE PROTEIN P2.
 FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
 FT CHAIN 432 447 CORE PROTEIN P1.
 FT CHAIN 448 511 CORE PROTEIN P6.
 FT ZN FING 389 406 CCHC-TYPE 1.
 FT ZN FING 410 427 CCHC-TYPE 2.
 FT LIPID 1 N-myristoyl glycine (in host) (By similarity).
 SQ SEQUENCE 511 AA; 57125 MW; 97C95188BE80D458 CRC64;
 Query Match 35.7%; Score 50; DB 1; Length 511;
 Best Local Similarity 78.6%; Pred. No. 6;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 15 TERGRSLYNTVATL 28
 Db 71 SEELRSLYNTVATL 84
 ID GAG_HV1B5 STANDARD; PRT; 511 AA.
 AC P04593;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE GAG polyprotein [Contains: Core proteins P17, P24, P2, P7, P1, P6].
 GN GAG.
 OS Human immunodeficiency virus type 1 (BHS isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11682;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
 RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
 RA Lautenberger J.A., Pappas T.S., Ghayab J., Chang N.T., Gallo R.C.,
 RA Wong-Staal F.;
 "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 Nature 313:277-284(1985).
 RT Nature 313:277-284(1985).
 RL Nature 313:277-284(1985).
 CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
 ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
 REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
 MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
 RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
 CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
 BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
 CC -!- PTM: The p24 protein is phosphorylated.
 CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
DR EMBL; K02012; AAA44652.1; -.
DR HSSP; P05888; 1AAF.
DR HIV; K02012; GAGSBS.
DR InterPro; IPR000721; Gag_P24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovlt_p17.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_P24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; CCHCZNFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR POSITE; PS50158; ZF_CCHC; 2.
DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; Lipoprotein.
FT INIT_MET 0
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P1.
FT CHAIN 448 511 CORE PROTEIN P6.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 N-myristoyl glycine (in host) (By
SEQUENCE 511 AA; 57107 MW; E0AE72D3CBE7A80 CRC64;
similarity).
Query Match 35.7%; Score 50; DB 1; Length 511;
Best local Similarity 78.6%; Pred. No. 6;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGSLYNTVATL 28
Db 71 SEELRSLYNTVATL 84

RESULT 15
GAG_HV1BR STANDARD; PRT; 511 AA.
AC P03348;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GAG polyprotein [contains: Core proteins P17, P24, P2, P7, P1, P6].
GN GAG.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retrovirda; Retroviridae; Lentivirus.
OC NCB1_TaxID=11686;
OX 11
RP SEQUENCE FROM N.A.
RX MEDLINE=85099333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
RN 12
RP REVISIONS TO 459-470.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -!- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE
ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL
REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM
MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY
RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY
BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.
CC -!- PTM: The P24 protein is phosphorylated.
CC -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02013; AAB59747.1; -.
DR HSSP; P05888; 1AAF.
DR HIV; K02013; GAGSBRU.
DR InterPro; IPR000721; Gag_P24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR000071; Retrovlt_p17.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_P24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; CCHCZNFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR POSITE; PS50158; ZF_CCHC; 2.
DR AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;
KW Zinc-finger; Repeat; Lipoprotein.
FT INIT_MET 0
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).
FT CHAIN 363 376 CORE PROTEIN P2.
FT CHAIN 377 431 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).
FT CHAIN 432 447 CORE PROTEIN P1.
FT CHAIN 448 511 CORE PROTEIN P6.
FT ZN_FING 389 406 CCHC-TYPE 1.
FT ZN_FING 410 427 CCHC-TYPE 2.
FT LIPID 1 N-myristoyl glycine (in host) (By
SEQUENCE 511 AA; 57095 MW; 55972701A8434673 CRC64;
similarity).
Query Match 35.7%; Score 50; DB 1; Length 511;
Best local Similarity 78.6%; Pred. No. 6;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 15 TERGSLYNTVATL 28
Db 71 SEELRSLYNTVATL 84
```

Search completed: July 20, 2004, 06:31:00
Job time : 7.5082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 28.2295 Seconds

(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166B-273

Sequence: 1 GROYIKANSKFIGITERGRSLVNTVATL 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	50.0	1310	2 Q93N27	Q93N27 clostridium
2	55	39.3	342	16 Q98FZ0	Q98FZ0 rhizobium
3	55	39.3	499	15 Q8AU75	Q8AU75 human
4	54	38.6	80	15 Q9YK29	Q9YK29 human
5	54	38.6	82	15 Q9WEV6	Q9WEV6 human
6	54	38.6	82	15 Q9YK77	Q9YK77 human
7	54	38.6	101	15 Q9INB0	Q9INB0 human
8	54	38.6	102	15 Q9INB3	Q9INB3 human
9	54	38.6	103	15 Q9IND2	Q9IND2 human
10	54	38.6	119	15 Q9J661	Q9J661 human
11	54	38.6	121	15 Q99583	Q99583 human
12	54	38.6	121	15 Q4I399	Q4I399 human
13	54	38.6	121	15 Q99580	Q99580 human
14	54	38.6	126	15 Q99551	Q99551 human
15	54	38.6	126	15 Q9WV66	Q9WV66 human
16	54	38.6	126	15 Q995R8	Q995R8 human

17	54	38.6	127	15 Q77209	Q77209 human
18	54	38.6	128	15 Q8J6D4	Q8J6D4 human
19	54	38.6	129	15 Q55392	Q55392 human
20	54	38.6	129	15 Q7S1Q7	Q7S1Q7 human
21	54	38.6	130	15 Q8Q7Q9	Q8Q7Q9 human
22	54	38.6	130	15 Q76947	Q76947 human
23	54	38.6	130	15 Q78602	Q78602 human
24	54	38.6	130	15 Q76945	Q76945 human
25	54	38.6	130	15 Q77179	Q77179 human
26	54	38.6	130	15 Q78604	Q78604 human
27	54	38.6	130	15 Q77176	Q77176 human
28	54	38.6	130	15 Q76940	Q76940 human
29	54	38.6	130	15 Q8Q7T3	Q8Q7T3 human
30	54	38.6	130	15 Q78605	Q78605 human
31	54	38.6	130	15 Q77174	Q77174 human
32	54	38.6	130	15 Q8Q7Q8	Q8Q7Q8 human
33	54	38.6	130	15 Q78599	Q78599 human
34	54	38.6	130	15 Q78600	Q78600 human
35	54	38.6	130	15 Q78601	Q78601 human
36	54	38.6	130	15 Q77177	Q77177 human
37	54	38.6	130	15 Q76939	Q76939 human
38	54	38.6	130	15 Q55384	Q55384 human
39	54	38.6	130	15 Q78611	Q78611 human
40	54	38.6	130	15 Q8Q7S5	Q8Q7S5 human
41	54	38.6	130	15 Q78608	Q78608 human
42	54	38.6	130	15 Q78613	Q78613 human
43	54	38.6	130	15 Q76942	Q76942 human
44	54	38.6	130	15 Q76943	Q76943 human
45	54	38.6	130	15 Q77210	Q77210 human

ALIGNMENTS

RESULT 1

Q93N27	PRELIMINARY;	PRT;	1310 AA.
AC Q93N27;			
DT 01-DEC-2001 (TREMBLrel. 19, Created)			
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE Tetanus toxin (Fragment).			
OS Clostridium tetani.			
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC Clostridium.			
OX NCBI_TaxID=1513;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Shunlin Z., Dianliang L.,			
RT "Cloning and sequence analysis of tetanus toxin gene,"			
RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.			
DR EMBL; AF389424; AA072964.2; -			
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.			
DR GO; GO:0008237; F:metallopeptidase activity; IEA.			
DR GO; GO:0015070; F:toxin activity; IEA.			
DR GO; GO:0008270; F:zinc ion binding; IEA.			
DR GO; GO:0009405; P:patogenesis; IEA.			
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR InterPro; IPR008985; Cona_like_1ec_g1.			
DR InterPro; IPR001064; Crystallin.			
DR InterPro; IPR002160; Kunitz_legume.			
DR InterPro; IPR000395; Peptidase_M27.			
DR InterPro; IPR006025; Pept_M_zn_BS.			
DR Pfam; PF01742; Peptidase_M27; 1.			
DR PRINTS; PR00760; BONTOTOXISIN.			
DR PRODOM; PD001963; Bontotoxin1.			
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.			
DR PROSITE; PS00142; ZINC_PROTEASE; 1.			
FT NON_TER			
FT 1			
FT 1310			
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418F450 CRC64;			

Query Match

50.0%; Score 70; DB 2; Length 1310;

Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QYIKANSKFIGITE 16
Db 831 QYIKANSKFIGITE 844

RESULT 2

Q98FZ0 PRELIMINARY; PRT; 342 AA.

AC Q98FZ0; 01-OCT-2001 (TRENBLrel. 18, Created)

DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)

DE Hypoetical protein ml13560.

GN ML13560.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OX Phyllobacteriaceae; Mesorhizobium.

NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAF303099;

MDLINE=21082930; PubMed=11214968;

RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Matanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL; AP003002; BAB50426.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 342 AA; 37582 MW; 39597D57D5D11A1 CRC64;

Query Match 39.3%; Score 55; DB 16; Length 342;
Best Local Similarity 41.7%; Pred. No. 4.6;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 3 QYIKANSKFIGITERGRSLYNTVA 26
Db 82 RLKAGSDFIGVADTGWFFCTVA 105

RESULT 3

Q8AU75 PRELIMINARY; PRT; 499 AA.

AC Q8AU75; 01-MAR-2003 (TRENBLrel. 23, Created)

DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)

DE Gag protein.

GN Gag.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=99ZASW7;

MDLINE=22190101; PubMed=12201911;

RA Papatianopoulos M.A., Cilliers T., Morris L., Mokili J.L.,

RA Dowling W., Bix D.U., McCutchan F.E.;

RT "Full-length Genome Analysis of HIV-1 Subtype C Utilizing CXCR4 and

Inter-subtype Recombinants Isolated in South Africa."

RL AIDS Res. Hum. Retroviruses 18:879-886(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=99ZASW7;

RA Papatianopoulos M.A., Morris L., Mokili J., Dowling W., Carr J.K.,

RA McCutchan F.;

Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF413966; AAN47118.1; -

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR GO; GO:0016032; P:vital life cycle; IEA.

DR InterPro; IPR000721; Gag_p24.

DR InterPro; IPR000071; Retrovir_p17.

DR InterPro; IPR008916; Retrov_capsid_C.

DR InterPro; IPR008919; Retrov_capsid_N.

DR InterPro; IPR001878; ZnF_CCHC.

DR Pfam; PF00540; Gag_p17; 1.

DR Pfam; PF00607; Gag_p24; 1.

DR Pfam; PF00098; zf-CCHC; 2.

DR PRINTS; PR00939; C2HCNPFINGER.

DR PRINTS; PR00234; HIVMATRIX.

DR SMART; SM00343; ZNF_CCHC; 2.

DR PROSITE; PS50158; ZF_CCHC; 2.

SQ SEQUENCE 499 AA; 55941 MW; 1B4F1D83189BD712 CRC64;

Query Match 39.3%; Score 55; DB 15; Length 499;
Best Local Similarity 85.7%; Pred. No. 7;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
Db 72 TEERVSLYNTVATL 85

RESULT 4

Q9YKZ9 PRELIMINARY; PRT; 80 AA.

AC Q9YKZ9; 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DE Gag protein (Gag polyprotein) (Fragment).

GN Gag.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DG5218;

MDLINE=99210134; PubMed=10195759;

RA Naghavi M.H., Salminen M.O., Sonnerborg A., Vahine A.;

RT "DNA sequence of the long terminal repeat of human immunodeficiency

RT virus type 1 subtype A through G."

RL AIDS Res. Hum. Retroviruses 15:485-488(1999).

DR EMBL; AF106178; AAD16923.1; -

DR GO; GO:0019012; C:viralon; IEA.

DR InterPro; IPR000071; Retrovir_p17.

DR Pfam; PF00540; Gag_p17; 1.

DR PRINTS; PR00234; HIVMATRIX.

KW AIDS; Core protein; Polypotein.

FT NON TER 1

FT NON TER 80

SQ SEQUENCE 80 AA; 9341 MW; E556DE35D7F1042F CRC64;

Query Match 38.6%; Score 54; DB 15; Length 80;
Best Local Similarity 85.7%; Pred. No. 1.3;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
Db 46 TEERVSLYNTVATL 59

RESULT 5
Q9WEV6 PRELIMINARY; PRT; 82 AA.

AC Q9WEV6; 01-NOV-1999 (TRENBLrel. 12, Created)

DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZMS5710;
 RX MEDLINE=99210134; PubMed=10195759;
 RA Naghavi M.H., Salminen M.O., Sommerborg A., Vahne A.;
 RT "DNA sequence of the long terminal repeat of human immunodeficiency
 virus type 1 subtype A through G.";
 RL AIDS Res. Hum. Retroviruses 15:485-488(1999).
 DR EMBL: AF106192; AAD16937.1; -
 DR GO: GO:0019012; C:viral; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR000071; Retroviral_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS: Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 9229 MW; 7F8CDF198904E579 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 82;
 Best Local Similarity 85.7%; Pred. No. 1.3;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 48 TEELRSLYNTVATL 61

RESULT 6
 Q9YKY7 PRELIMINARY; PRT; 82 AA.
 ID Q9YKY7
 AC Q9YKY7
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-OCT-2003 (TrEMBLrel. 15, Last sequence update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SO6173;
 RX MEDLINE=99210134; PubMed=10195759;
 RA Naghavi M.H., Salminen M.O., Sommerborg A., Vahne A.;
 RT "DNA sequence of the long terminal repeat of human immunodeficiency
 virus type 1 subtype A through G.";
 RL AIDS Res. Hum. Retroviruses 15:485-488(1999).
 DR EMBL: AF106193; AAD16938.1; -
 DR GO: GO:0019012; C:viral; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR000071; Retroviral_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS: Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 9475 MW; 66A40A2AD4B4A30 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 82;
 Best Local Similarity 85.7%; Pred. No. 1.3;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 48 TEELRSLYNTVATL 61

RESULT 7
 Q9INE0 PRELIMINARY; PRT; 101 AA.
 ID Q9INE0
 AC Q9INE0
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DRCM17;
 RX MEDLINE=99260287; PubMed=10331444;
 RA Mokili J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,
 RA Green S.D., Peutherer J.F., Simmonds P.;
 RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural
 Democratic Republic of Congo.";
 RL AIDS Res. Hum. Retroviruses 15:655-664(1999).
 DR EMBL: AF144809; AAF69035.1; -
 DR GO: GO:0019012; C:viral; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR000071; Retroviral_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS: Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 101
 SQ SEQUENCE 101 AA; 11256 MW; 3B601426CF7841C9 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 101;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 33 TEELRSLYNTVATL 46

RESULT 8
 Q9INA3 PRELIMINARY; PRT; 102 AA.
 ID Q9INA3
 AC Q9INA3
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DRCM17;
 RX MEDLINE=99260287; PubMed=10331444;
 RA Mokili J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,
 RA Green S.D., Peutherer J.F., Simmonds P.;
 RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural
 Democratic Republic of Congo.";
 RL AIDS Res. Hum. Retroviruses 15:655-664(1999).
 DR EMBL: AF144846; AAF69072.1; -
 DR GO: GO:0019012; C:viral; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR000071; Retroviral_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS: Core protein; Polypeptide.
 FT NON_TER 1
 FT NON_TER 102
 SQ SEQUENCE 102 AA; 11423 MW; 69A9761220982C5E CRC64;

Query Match 38.6%; Score 54; DB 15; Length 102;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 33 TEELRSLYNTVATL 46

RESULT 9
 Q9IND2 PRELIMINARY; PRT; 103 AA.

AC Q9IND2;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.

OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 NCBI_TaxID=11676;
 [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DCOM13;
 RX MEDLINE=9260287; PubMed=1031444;
 RA McKill J.L., Wade C.M., Burns S.M., Cutting W.A., Bopopi J.M.,
 Green S.D., Penhagher J.F., Simmonds P.;
 RT "Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural
 Democratic Republic of Congo";
 RL AIDS Res. Hum. Retroviruses 15:655-664(1999).
 DR EMBL; AF14817; AAF69043.1; -;
 DR GO; GO:0019012; C:virion; IEA.
 DR InterPro; IPR000071; Retrovir_p17.
 DR Pfam; PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 DR AIDS; Core protein; Polypeptide.
 KW NON_TER 1
 FT NON_TER 103
 SQ SEQUENCE 103 AA; 11577 MW; AAD469474ED731E2 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 33 TEELRSLYNTVATL 46

RESULT 10
 Q9J661 PRELIMINARY; PRT; 119 AA.

AC Q9J661;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.

OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 NCBI_TaxID=11676;
 [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DK96AIP41;
 RA Iversen A.K., Learn G.H., Fugger L., Gerstoft J., Mullins J.I.,
 Skirboj P.;
 RT "Presence of multiple HIV subtypes and a high frequency of subtype
 RT chimeric viruses in heterosexually infected women";
 RL J. Acquir. Immune Defic. Syndr. Hum. Retroviro. 22:325-332(1999).
 DR EMBL; AF191484; AAF37550.1; -;
 DR GO; GO:0019012; C:virion; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000071; Retrovir_p17.
 DR Pfam; PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 DR AIDS; Core protein; Polypeptide.
 KW NON_TER 1
 FT NON_TER 119
 SQ SEQUENCE 119 AA; 13536 MW; F1366864F6F28DD9 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 119;
 Best Local Similarity 85.7%; Pred. No. 2;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 42 TEELRSLYNTVATL 55

RESULT 11
 Q995R9 PRELIMINARY; PRT; 121 AA.

AC Q995R9;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE Gag protein (Gag polypeptide) (Fragment).
 GN GAG.

OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 NCBI_TaxID=11676;
 [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Y9;
 RX Vitrelli D.L.;
 RT "Immunological and virological characteristics of HIV infected
 RT individuals in different risk groups in Madras, India";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF250270; AAK34984.1; -;
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000071; Retrovir_p17.
 DR Pfam; PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 DR AIDS; Core protein; Polypeptide.
 KW NON_TER 1
 FT NON_TER 121
 SQ SEQUENCE 121 AA; 13968 MW; 369B37CDA213AD4F CRC64;

Query Match 38.6%; Score 54; DB 15; Length 121;
 Best Local Similarity 85.7%; Pred. No. 2.1;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 15 TERGRSLYNTVATL 28
 |||||
 DB 48 TEELRSLYNTVATL 61

RESULT 12
 Q41399 PRELIMINARY; PRT; 121 AA.

AC Q41399;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, last annotation update)
 DE P17 (Gag polypeptide) (Fragment).
 GN GAG.

OS Human immunodeficiency virus 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 NCBI_TaxID=11676;
 [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=ET2555;
 RA Shereta K., Johansson B., Sommerborg A.;

RT "Distribution of HIV-1 subtypes in Ethiopia."
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBD databases
 DR EMBL, U56345; AAB67649.1; -
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR Interpro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRX.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 121 aa; 121
 SQ SEQUENCE 121 AA; 13849 MW; 89661A1FE09CDB3 CRC64;

Query Match 38.6%; Score 54; DB 15; Length 121;
Best Local Similarity 85.7%; Pred. No. 2.1;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	15	TERGRSLYNTVATL	28
Db	72	TEELRSLYNTVATL	85

RESULT 13			
099550			
ID	Q99550	PRELIMINARY;	PRT; 124 AA.
AC	Q99550;		
DT	01-JUN-2001	(TREMBLrel. 17, Created)	
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)	
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)	
DE	Gag protein (Gag polypeptide) (Fragment).		
GN	GAG.		
OS	Human immunodeficiency virus 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_Taxid=11676;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=Y8;		
RA	Virrell D.L.;		
RT	"Immunological and virological characteristics of HIV infected individuals in different risk groups in Madras, India.,"		
RL	Submitted (MAR-2000) to the EMBL/Genbank/DBD databases.		
DR	EMBL; AF050269; AAC34983.1; -		
DR	GO; GO:0019012; C:virion; IEA.		
DR	GO; GO:0005198; F:structural molecule activity; IEA.		
DR	InterPro; IPR000071; Retrovix_p17.		
DR	Pfam; PF00540; Gag_p17; 1.		
DR	PRINTS; PR00234; HIVIMATRX.		
KW	AIDS; Core protein; Polyprotein.		
FT	NON_TER	1	
FT	NON_TER	124	124
SQ	SEQUENCE	124 AA;	14178 MW: CSBIB05020568A77 CXC64;

Query Match	38.6%;	Score 54;	DB 15;	Length 124;
Best Local Similarity	85.7%;	Pred. No. 2.1;		
Matches 12;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	15	TERGRSLYNTVATL	28
Db	46	TEELRSLYNTVATL	59

Accession	Protein Name	Length (AA)
Q995S1	PRELIMINARY; PRT;	126 AA.
Q995S1	Q995S1;	
DDT	01-JUN-2001 (TREMBLrel. 17, Created)	
DDT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
DDT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
DEE	Gag protein (Gag polypeptide) (Fragment).	
DDT	GAG	
DDT	Human immunodeficiency virus 1.	
DDT	Viruses; Retroviral viruses; Retroviridae; Lentivirus	
DDT	NCBI_TaxID=11676;	
DDT	[1]	

RP SEQUENCE FROM N.A.
PC STPATN-Y7.

RC STRAIN=Y7;
PA W/mo17 D r

Yirrell D.L.

RT "Immunological and virological characteristics of HIV infected

Rt "Immunological and virological characteristics of HIV in individuals in different risk groups in Madras, India"

RI	individuals in different risk groups
RL	Submitted (MAR-2000) to the EMRL/CO

Submitted (MAR-2000) to the EMBL
EMBL: AF250268; AAK34987 1.

DR	GO:0019012; C:virion; IEA.
DR	GO:00005198; F:structural molecule activity; IEA.
DR	InterPro: IPR000071; Retrovir_p17.
DR	Pfam: PF00540; Gag_P17: 1
DR	PRINTS: PR00234; HIVMATRIX.
KW	AIDS; Core protein; Polyprotein.
FT	NON TER
FT	1
FT	126
SEQUENCE	126 AA; 14336 MW; SC561821D8ED0DA4 CRC64;

	Query March	Best Local Similarity	Score 54;	DB 15;	Length 126;
		85.7%;	Pred. No. 2.2;		
Matches	12;	Conservative	0;	Mismatches	2; Indels 0; Gaps 0;
Oy	15 TERCGRSYNTVAII	28 			
Db	48 TEELRSLYNTVAII	61 			

RESULT 15	
Q9WEV6	
Q9WEV6	
Q9WEV6	
Q9WEV6	
PRELIMINARY:	
PRT:	126 AA.

NOT FOR PUBLICATION

DT 01-NOV-1999 (TREMBLRel. 12, Created

DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)
DT	01-NOV-1999	(TREMBLrel. 12, Created)

D1	01-NOV-1999	(TREMBlrel. 12, Last sec
DT	01-JUN-2003	(TREMBlrel. 24, Last am

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gag protein (Gag polypeptide) (Fragment)

DE	Gag protein (Gag polypeptide)
GN	GAG.

GN	GAG.
OS	Human immunodeficiency virus 1

OC Viruses, Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]_Taxid=11676;
 RP SEQUENCE FROM N.A.
 RC STRAIN=837;
 RX MEDLINE=99181514; PubMed=10081673;
 RA Ramos A., Tanuri A., Schechter M., Rayfield M.A., Hu D.-J.,
 Rabril M.C., Banea C.I., Baggs J., Pieniazek D.,
 RT "Dual and recombinant infections: an integral part of the HIV-1

RL	Emerging Infect. Dis.	5:65-74(1999).
Pravetani	Influenza A Virus	-
DR	EMBL	AFL15451; RAD21596.1; -
DR	GO	GO:0019012 C:virion; IEA.
DR	GO	GO:0005198; F:structural molecule activity; IEA.
DR	InterPro	IPR000071; Retroviral_P17.
DR	Pfam	PF00540; Gag_p17; 1.
DR	PRINTS	PR00234; HIVMAMRIX.
ALDS	AIDS	Core protein; Polyprotein.
FT	NON_TER	1 126
FT	NON_TER	1 126
ISQ	SEQUENCE	126 AA; 14137 MW; 6EC9E47F2FD3CEA0 CRC64;
 Query Match		
Best Local Similarity 38.6%; Score 54; DB 15; Length 126;		
Matches 12; Conservative 85.7%; Pred. No. 2.2;		
Indels 0; Mismatches 2; Gaps 0;		
 DB		
15	TERGRSLNTVATL	28
72	TEELRSLSIVTAVTL	85

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 41.7705 Seconds

(without alignments)
189,400 Million cell updates/sec

Title: US-09-673-166b-273

Sequence: 140

Sequence: 1 GROYIKANKKFGITGRGSLNYVTATL 28

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	70.0	28	2	AAV53552
2	88.5	63.2	25	4	AAAB84738
3	87	62.1	29	2	AAV53551
4	83	59.3	29	2	AAW48991
5	83	59.3	29	2	AAW48992
6	83	59.3	29	2	AAW09326
7	83	59.3	29	2	AAW09327
8	83	59.3	29	3	AAAB12418
9	83	59.3	29	3	AAAB12419
10	83	59.3	29	4	AAE05438
11	83	59.3	29	4	AAE05437
12	83	59.3	29	5	AAU10851
13	83	59.3	29	5	AAU10850
14	82	58.6	28	5	AAU11422
15	82	57.9	29	5	ABR79182
16	80	57.1	158	5	AAW81330
17	80	57.1	158	5	ABR07275
18	79	56.4	218	3	AAV49253
19	79	56.4	240	3	AAV49254
20	79	56.4	390	3	AAV49255
21	79	56.4	390	6	AAO30491
22	79	56.4	514	6	AAO30490
23	79	56.4	514	6	AAO30495
24	79	56.4	517	6	AAO30492
25	77	55.0	750	3	AAV92645

26	77	55.0	750	3	AAV92638	AAV92638 Mutant hu
27	77	55.0	750	3	AAV92640	AAV92640 Mutant hu
28	76	54.3	25	3	AAV92651	AAV92651 psmep008
29	76	54.3	29	2	AAAB83561	AAAB83561 Ige CH4 I
30	76	54.3	37	2	AAAB65389	AAAB65389 Universal
31	76	54.3	37	2	AAAB65383	AAAB65383 Universal
32	76	54.3	47	2	AAAB62723	AAAB62723 LHRH-cont
33	76	54.3	750	3	AAV92627	AAV92627 Mutant hu
34	75.5	53.9	750	3	AAV92639	AAV92639 Mutant hu
35	75	53.6	25	3	AAV92652	AAV92652 psmep009
36	75	53.6	50	2	AAW06132	AAW06132 Anti-cho
37	75	53.6	50	2	AAW46447	AAW46447 CRTP B ce
38	75	53.6	109	4	AAAB20145	AAAB20145 Growth di
39	75	53.6	124	3	AAAB45492	AAAB45492 Modified
40	75	53.6	124	3	AAAB45517	AAAB45517 Modified
41	75	53.6	158	2	AAAB81329	AAAB81329 TNF2-4, a
42	75	53.6	158	5	ABR07281	ABR07281 Human TNF
43	75	53.6	216	3	AAV92665	AAV92665 MGC-1 ana
44	75	53.6	693	3	AAV92649	AAV92649 Mutant hu
45	75	53.6	750	3	AAV92630	AAV92630 Mutant hu

ALIGNMENTS

RESULT 1	
AAV53552	
ID	AAV53552 standard; protein; 28 AA.
XX	
AC	AAV53552;
XX	
DT	18-JAN-2000 (first entry)
XX	
DE	Lipopeptide #3.
XX	
KM	Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW	electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KW	human immunodeficiency virus; hepatitis B virus; papilloma virus;
XX	melanoma; malaria; parasite.
OS	Synthetic.
OS	Homo sapiens.
XX	
FH	Key
FT	Modified-site
FT	1 /note= "contains palmitoyl residue attached at the N-
FT	terminus and on the epsilon carbon of the side chain"
XX	
FN	FR2776926-A1.
XX	
PD	08-OCT-1999.
XX	
PF	07-APR-1998; 98FR-00004323.
XX	
PR	07-APR-1998; 98FR-00004323.
XX	
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA	(CNRS) CNRS CENT NAT RECH SCI.
PA	(INSP) INST PASTEUR ILLRE.
XX	
PI	Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
XX	
DR	WPI; 1999-583113/50.
XX	
PT	New lipopeptide containing lipid regions and two epitopes, all separated
PT	by peptide spacers that impart hydrophilicity, useful in vaccines.
XX	
PS	Example 1; Page 9; 35pp; French.
XX	
CC	The invention relates to the generation of a lipopeptide comprising at
CC	least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC	epitope and at least one lipid residue with (i) the epitopes and lipid
CC	portion and (ii) the epitopes, being separated independently by peptide

CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAY5301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 28 AA;

Query Match 70.0%; Score 98; DB 2; Length 28;
Best Local Similarity 95.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GROYIKANSKFIGITERGRS 20
1 GROYIKANSKFIGITERGRA 20

RESULT 2

ID AAB84738 standard; peptide; 25 AA.

AC AAB84738;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of lipopeptide Nh2-K.

KW Lipid-tailed protein; mucosal membrane; immune system; lipoprotein;
KW B cell response; T cell response; intranasal immunisation;
KW sublingual immunisation.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "amidated residue with palmitoyl lipid residue
FT with 16 carbon atoms attached via an epsilon bond"

PN W0200141797-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000MO-EP012794.

PR 09-DEC-1999; 99US-0169952P.

PA (INSP) INST PASTERUR.

PA (UYLI-) UNIV LITTLE 2.

PI Druilhe P, Gras-Masse H, Benmohamed L,

DR WPI; 2001-451597/48.

PT Inducing systemic immune response comprises mucosal administration of

PS Claim 22; Page 26; 34pp; English.

CC The present sequence represents a lipid-tailed protein. The lipid moiety
CC is able to cross the mucosal membrane, and deliver an antigen to the
CC immune system. The lipoprotein is administered to a mucosal membrane to
CC induce an immune response. The lipoprotein induces a B cell and/or a T
CC cell response. The lipoprotein is used to provide intranasal or
CC sublingual immunization

SQ Sequence 25 AA;

Query Match 63.2%; Score 88.5; DB 4; Length 25;

Best Local Similarity 90.9%; Pred. No. 4.5e-07;
Matches 20; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 GROYI-KANSKFIGITERGRSL 21
1 GROYIKANSKFIGITERGRIL 22

RESULT 3
ID AAY53551 standard; protein; 29 AA.

AC AAY53551;

DT 18-JAN-2000 (first entry)

DE Lipopeptide #2.

KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
KW melanoma; malaria; parasite.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "contains palmitoyl residue attached at the N-
FT terminus and on the epsilon carbon of the side chain"

PN FR2776926-A1.

PD 08-OCT-1999.

PF 07-APR-1998; 98FR-00004323.

PR 07-APR-1998; 98FR-00004323.

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (CNRS) CNRS CENT NAT RECH SCI.

PA (INSP) INST PASTERUR LITTLE.

PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;

DR WPI; 1999-583113/50.

PT New lipopeptide containing lipid regions and two epitopes, all separated
PT by peptide spacers that impart hydrophilicity, useful in vaccines.

PS Example 1; Page 9; 35pp; French.

CC The invention relates to the generation of a lipopeptide comprising at
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC epitope and at least one lipid residue with (i) the epitopes and lipid
CC portion and (ii) the epitopes, being separated independently by peptide
CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAY5301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 29 AA;

Query Match 62.1%; Score 87; DB 2; Length 29;
Best Local Similarity 94.4%; Pred. No. 9.4e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 OYIKANSKFIGTTERGRS 20
 |||||
 4 OYIKANSKFIGTTERGRA 21

RESULT 4
 AAM48991
 ID AAM48991 standard; peptide; 29 AA.
 XX AAM48991;
 DT 23-SEP-1998 (first entry)
 XX
 DE Lipidated vaccine 1 against human cytomegalovirus infection.

XX Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; CD8+;
 KW cytotoxic T-lymphocyte; MHC; major histocompatibility complex;
 KM helper T-lymphocyte; HTL; CTL; palmitic acid.
 XX
 OS Synthetic.

FT Key Location/Qualifiers
 FT Modified-site 1
 FT Region /note= "attached to two palmitic acid molecules"
 FT 2..17
 FT Region /note= "Helper T-cell epitope"
 FT 18..20
 FT /note= "Spacer"
 FT 21..29
 FT Region /note= "Cytotoxic T-cell epitope referred to as the 495
 FT peptide"
 XX
 PN WO9821233-A2.
 XX
 PD 22-MAY-1998.
 XX
 PF 12-NOV-1997; 97WO-US020236.
 XX
 PR 12-NOV-1996; 96US-00747488.
 PR 14-OCT-1997; 97US-00950064.
 XX
 PA (CITY) CITY OF HOPE.
 XX
 PI Diamond DJ, York U;
 XX
 DR WPI; 1998-297862/26.
 XX
 PT Immunoreactive human cytomegalovirus epitopes - useful to vaccinate
 PT against infection.
 XX
 PS Example 11; Page 31; 56pp; English.
 XX

CC The present lipidated vaccine comprises of two palmitic acid molecules
 CC attached to a peptide portion. The peptide portion contains a helper T-
 CC lymphocyte (HTL) epitope attached to a spacer which in turn is attached
 CC to a cytotoxic T-lymphocyte (CTL) epitope. The CTL epitope is referred to
 CC as the 495 peptide as it represents residues 495-503 of pp65 from human
 CC cytomegalovirus (HCMV) laboratory strains AD169 and Towne. The invention
 CC provides peptides which are immunogenic epitopes that are recognised by
 CC the CD8+ class I major histocompatibility complex (MHC) restricted
 CC cytotoxic T-lymphocytes of patients harboring latent HCMV infection. The
 CC peptides are claimed to be capable of activating cytotoxic T-cell
 CC lymphocytes in the absence of active viral replication, and thus are
 CC useful for eliciting a cellular immune response against HCMV by normal
 CC and immunodeficient subjects. The immunological peptides can also be
 CC formulated as vaccines, such as the present one, which are claimed to be
 CC useful for protecting against HCMV infection, augmenting the immune
 CC system response to a HCMV infection or protecting against reactivation of
 CC a latent HCMV infection
 CC
 XX Sequence 29 AA;
 SO

Query Match 59.3%; Score 83; DB 2; Length 29;

Best Local Similarity 69.2%; Pred. No. 4.2e-06;
 Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 3 OYIKANSKFIGTTERGRSLYNATL 28
 |||||
 4 OYIKANSKFIGTTEAANLVPMVATV 29

RESULT 5
 AAM48992
 ID AAM48992 standard; peptide; 29 AA.
 XX AAM48992;
 DT 23-SEP-1998 (first entry)
 XX
 DE Lipidated vaccine 2 against human cytomegalovirus infection.

XX Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; CD8+;
 KW cytotoxic T-lymphocyte; MHC; major histocompatibility complex;
 KM helper T-lymphocyte; HTL; CTL; palmitic acid.
 XX
 OS Synthetic.

FT Key Location/Qualifiers
 FT Modified-site 1
 FT Region /note= "attached to three palmitic acid molecules"
 FT 4..17
 FT Region /note= "Helper T-cell epitope"
 FT 18..20
 FT /note= "Spacer"
 FT 21..29
 FT Region /note= "Cytotoxic T-cell epitope referred to as the 495
 FT peptide"
 XX
 PN WO9821233-A2.
 XX
 PD 22-MAY-1998.
 XX
 PF 12-NOV-1997; 97WO-US020236.
 XX
 PR 12-NOV-1996; 96US-00747488.
 PR 14-OCT-1997; 97US-00950064.
 XX
 PA (CITY) CITY OF HOPE.
 XX
 PI Diamond DJ, York U;
 XX
 DR WPI; 1998-297862/26.
 XX
 PT Immunoreactive human cytomegalovirus epitopes - useful to vaccinate
 PT against infection.
 XX
 PS Example 11; Page 31; 56pp; English.
 XX

CC The present lipidated vaccine comprises of three palmitic acid molecules
 CC attached to a peptide portion. The peptide portion contains a helper T-
 CC lymphocyte (HTL) epitope attached to a spacer which in turn is attached
 CC to a cytotoxic T-lymphocyte (CTL) epitope. The CTL epitope is referred to
 CC as the 495 peptide as it represents residues 495-503 of pp65 from human
 CC cytomegalovirus (HCMV) laboratory strains AD169 and Towne. The invention
 CC provides peptides which are immunogenic epitopes that are recognised by
 CC the CD8+ class I major histocompatibility complex (MHC) restricted
 CC cytotoxic T-lymphocytes of patients harboring latent HCMV infection. The
 CC peptides are claimed to be capable of activating cytotoxic T-cell
 CC lymphocytes in the absence of active viral replication, and thus are
 CC useful for eliciting a cellular immune response against HCMV by normal
 CC and immunodeficient subjects. The immunological peptides can also be
 CC formulated as vaccines, such as the present one, which are claimed to be
 CC useful for protecting against HCMV infection, augmenting the immune
 CC system response to a HCMV infection or protecting against reactivation of
 CC a latent HCMV infection
 CC
 XX Sequence 29 AA;
 SO

Sequence 29 AA; 59.3%; Score 83; DB 2; Length 29;
Query Match Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QYIKANSKFIIGITERGRSLYNTVATL 28
DB 4 QYIKANSKFIIGITEAANLVPMVATV 29

RESULT 6
AA09326
ID AA09326 standard; peptide; 29 AA.
AC AA09326;
XX
XX
DT 08-JUL-1999 (first entry)
DE Lipidated vaccine peptide SEQ ID NO:13.
XX
XX Human cytomegalovirus; hCMV; immunologically active peptide; vaccine;
KM immune response; cytotoxic T lymphocyte; CTL; immunostimulation;
KM infection; immunosuppression; bone marrow transplant; solid organ; heart;
KM AIDS.
XX
XX Synthetic.
OS
XX WO9919349-A1.
XX
XX 22-APR-1999.
PD
XX 11-MAY-1998; 98WO-US009652.
PF
XX 14-OCT-1997; 97US-00950064.
PR 10-FEB-1998; 98US-00021298.
XX
XX (CITY) CITY OF HOPE.
PI
XX Diamond DJ, York J;
PI
XX MPI; 1999-277590/23.
DR
XX
XX Immunogenic peptide cytotoxic T lymphocyte epitopes of human
PT cytomegalovirus.
PT
XX
XX Example 11; Page 32; 64pp; English.
PS
XX The present invention describes immunologically active peptides (IAPs)
CC capable of eliciting a cellular immune response to human cytomegalovirus
CC (hCMV). The IAP can be used in a (cellular) vaccine to augment the immune
CC system response to hCMV, or to provide immunity against hCMV. The IAP
CC (cellular) vaccine can also protect an individual having a latent hCMV
CC infection from reactivation. An antigen presenting cell can be used to
CC determine the presence or absence of hCMV-infected T lymphocytes. The
CC viral vector containing IAP encoding DNA can also be used to provide
CC immunity against hCMV. The IAP can be used to prepare hCMV-reactive human
CC cytotoxic T lymphocytes. The antigen presenting cells primed with the
CC IAPs can be used as diagnostic reagents to detect immunostimulation by
CC hCMV. They can also detect active hCMV infection or exposure to hCMV.
CC hCMV can cause opportunistic infections resulting in a variety of
CC complications in, e.g. immunosuppressed patients. The IAP vaccines impart
CC immunity to bone marrow transplant recipients, solid organ recipients,
CC heart patients, AIDS patients or women of child-bearing years, without
CC the need for ex vivo expansion of hCMV-specific cytotoxic T lymphocytes
CC (which requires sophisticated laboratory setup and is highly labor-
CC intensive and costly). The present sequence represents a lipidated
CC vaccine peptide from the present invention
XX
XX Sequence 29 AA;

Query Match 59.3%; Score 83; DB 2; Length 29;
Best Local Similarity 69.2%; Pred. No. 4.2e-06;

Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QYIKANSKFIIGITERGRSLYNTVATL 28
DB 4 QYIKANSKFIIGITEAANLVPMVATV 29

RESULT 7
AA09327
ID AA09327 standard; peptide; 29 AA.
AC AA09327;
XX
XX
DT 08-JUL-1999 (first entry)
DE Lipidated vaccine peptide SEQ ID NO:14.
XX
XX Human cytomegalovirus; hCMV; immunologically active peptide; vaccine;
KM immune response; cytotoxic T lymphocyte; CTL; immunostimulation;
KM infection; immunosuppression; bone marrow transplant; solid organ; heart;
KM AIDS.
XX
XX Synthetic.
OS
XX WO9919349-A1.
XX
XX 22-APR-1999.
PD
XX 11-MAY-1998; 98WO-US009652.
PF
XX 14-OCT-1997; 97US-00950064.
PR 10-FEB-1998; 98US-00021298.
XX
XX (CITY) CITY OF HOPE.
PI
XX Diamond DJ, York J;
PI
XX MPI; 1999-277590/23.
DR
XX
XX Immunogenic peptide cytotoxic T lymphocyte epitopes of human
PT cytomegalovirus.
PT
XX
XX Example 11; Page 32; 64pp; English.
PS
XX The present invention describes immunologically active peptides (IAPs)
CC capable of eliciting a cellular immune response to human cytomegalovirus
CC (hCMV). The IAP can be used in a (cellular) vaccine to augment the immune
CC system response to hCMV, or to provide immunity against hCMV. The IAP
CC (cellular) vaccine can also protect an individual having a latent hCMV
CC infection from reactivation. An antigen presenting cell can be used to
CC determine the presence or absence of hCMV-infected T lymphocytes. The
CC viral vector containing IAP encoding DNA can also be used to provide
CC immunity against hCMV. The IAP can be used to prepare hCMV-reactive human
CC cytotoxic T lymphocytes. The antigen presenting cells primed with the
CC IAPs can be used as diagnostic reagents to detect immunostimulation by
CC hCMV. They can also detect active hCMV infection or exposure to hCMV.
CC hCMV can cause opportunistic infections resulting in a variety of
CC complications in, e.g. immunosuppressed patients. The IAP vaccines impart
CC immunity to bone marrow transplant recipients, solid organ recipients,
CC heart patients, AIDS patients or women of child-bearing years, without
CC the need for ex vivo expansion of hCMV-specific cytotoxic T lymphocytes
CC (which requires sophisticated laboratory setup and is highly labor-
CC intensive and costly). The present sequence represents a lipidated
CC vaccine peptide from the present invention
XX
XX Sequence 29 AA;

Query Match 59.3%; Score 83; DB 2; Length 29;
Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

DB 4 QYIKANSKFIGITEAANLVPMVATV 29

RESULT 8

AAB12418

ID AAB12418 standard; peptide; 29 AA.

AC AAB12418;

DT 20-OCT-2000 (first entry)

DE HCMV lipidated vaccine peptide SEQ ID NO:13.

XX Immunogenic; cytotoxic T lymphocytes; CTL; cytomegalovirus; HCMV;
XX infection; immune response; vaccine; immunostimulant; antiviral;
XX immunosuppressive; immunity; immunisation.

OS Human herpesvirus 5.

FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminally attached to (Pam)2 where Pam is
FT palmitic acid"

XX US6074645-A.

XX 13-JUN-2000.

XX 11-MAY-1998; 98US-00075257.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX (CITY) CITY OF HOPE.

XX York J, Diamond DJ;

XX WPI; 2000-430383/37.

PT New immunogenic peptides useful for providing immunity against
PT cytomegalovirus (CMV) infections, are capable of eliciting cellular
PT immune response to human CMV.

XX Example 11; Col 15; 18pp; English.

XX The present invention describes a cellular vaccine (I), which elicits a
XX MHC Class I cellular immune response to human cytomegalovirus (HCMV), and
XX comprises a peptide selected from (i) - (v), provided that (i) is not
XX another peptide (vi): (i) Asn Xaa1 Val Pro Met Val Ala Thr Xaa2; (ii) Tyr
XX Ala Leu Xaa5; (iv) Thr Pro Arg Val Thr Gly Gly Ala Xaa6; and (v) Phe
XX Pro Thr Lys Asp Val Ala Leu, where: Xaa1 = Ser, Thr or Leu; Xaa4 = Val
XX or Thr; Xaa5 = Leu, Arg or Lys; Xaa6 = Leu, Ile, Met, Thr or Val;
XX Asn Leu Val Pro Met Val Ala Thr Val (vi). The peptides are useful for
XX providing immunity against future infections by HCMV, and for augmenting
XX the immune response of an individual who is latently infected with HCMV
XX and is at risk for reactivation of HCMV infection. The peptides are also
XX useful for imparting immunity to a bone marrow transplant recipient, a
XX solid organ recipient, a heart patient, an AIDS patient or a woman of
XX child-bearing years, without the need for ex vivo expansion of HCMV-
XX specific CTL. These peptides can directly stimulate cytotoxic T
XX lymphocytes (CTL) in vitro, thus can be used in an assay to determine the
XX degree of immunostimulation caused by HCMV. These may also be used to
XX distinguish individuals who are seropositive from those who have not been
XX exposed to HCMV and in the study of the Class I antigen-processing
XX pathway for HCMV proteins. The present sequence represents a peptide used
XX in the exemplification of the present invention

SQ Sequence 29 AA;

Query Match

59.3%; Score 83; DB 3; Length 29;

Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QYIKANSKFIGITEGRSLYNTVATL 28

DB 4 QYIKANSKFIGITEAANLVPMVATV 29

RESULT 9

AAB12419

ID AAB12419 standard; peptide; 29 AA.

AC AAB12419;

DT 20-OCT-2000 (first entry)

DE HCMV lipidated vaccine peptide SEQ ID NO:14.

XX Immunogenic; cytotoxic T lymphocytes; CTL; cytomegalovirus; HCMV;
XX infection; immune response; vaccine; immunostimulant; antiviral;
XX immunosuppressive; immunity; immunisation.

OS Human herpesvirus 5.

FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminally attached to (Pam)3 where Pam is
FT palmitic acid"

XX US6074645-A.

XX 13-JUN-2000.

XX 11-MAY-1998; 98US-00075257.

XX 12-NOV-1996; 96US-00747488.

XX 14-OCT-1997; 97US-00950064.

XX 10-FEB-1998; 98US-00021298.

XX (CITY) CITY OF HOPE.

XX York J, Diamond DJ;

XX WPI; 2000-430383/37.

PT New immunogenic peptides useful for providing immunity against
PT cytomegalovirus (CMV) infections, are capable of eliciting cellular
PT immune response to human CMV.

XX Example 11; Col 15; 18pp; English.

XX The present invention describes a cellular vaccine (I), which elicits a
XX MHC Class I cellular immune response to human cytomegalovirus (HCMV), and
XX comprises a peptide selected from (i) - (v), provided that (i) is not
XX another peptide (vi): (i) Asn Xaa1 Val Pro Met Val Ala Thr Xaa2; (ii) Tyr
XX Ala Leu Xaa5; (iv) Thr Pro Arg Val Thr Gly Gly Ala Xaa6; and (v) Phe
XX Pro Thr Lys Asp Val Ala Leu, where: Xaa1 = Ser, Thr or Leu; Xaa4 = Val
XX or Thr; Xaa5 = Leu, Arg or Lys; Xaa6 = Leu, Ile, Met, Thr or Val;
XX Asn Leu Val Pro Met Val Ala Thr Val (vi). The peptides are useful for
XX providing immunity against future infections by HCMV, and for augmenting
XX the immune response of an individual who is latently infected with HCMV
XX and is at risk for reactivation of HCMV infection. The peptides are also
XX useful for imparting immunity to a bone marrow transplant recipient, a
XX solid organ recipient, a heart patient, an AIDS patient or a woman of
XX child-bearing years, without the need for ex vivo expansion of HCMV-
XX specific CTL. These peptides can directly stimulate cytotoxic T
XX lymphocytes (CTL) in vitro, thus can be used in an assay to determine the
XX degree of immunostimulation caused by HCMV. These may also be used to
XX distinguish individuals who are seropositive from those who have not been
XX exposed to HCMV and in the study of the Class I antigen-processing
XX pathway for HCMV proteins. The present sequence represents a peptide used

CC	in the exemplification of the present invention
XX	
SQ	Sequence 29 AA;
	Query Match 59.3%; Score 83; DB 3; Length 29;
	Best local Similarity 69.2%; Pred. No. 4.2e-06;
	Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
OY	3 QYIKANSKFIGITERGRLVNTATL 28 : 4 QYIKANSKFIGITEAANLVNMTATV 29
Db	
	RESULT 10
AAE05438	ID AAE05438 standard; peptide: 29 AA.
XX	AAE05438;
AC	
XX	24-SEP-2001 (first entry)
DT	
XX	
DE	Human cytomegalovirus (HCMV) vaccine peptide #2.
XX	
KM	Human cytomegalovirus; HCMV; infection; vaccine; cytotoxic T lymphocyte;
XX	CTL; CTL precursor; CTLp; Class I antigen;
KW	major histocompatibility complex; MHC; cellular immune response;
KW	CTL activator; CTLp activator; immunostimulant.
XX	
OS	Human cytomegalovirus.
XX	
PH	Key Location/Qualifiers
FT	1 /note= "Cysteine modified with palmitic acid"
FT	4..17
FT	/label= HTL_epitope
FT	18..20
FT	/note= "amino acid spacer"
FT	21..29
FT	/label= CTL_epitope
XX	
PN	US6251399-B1.
XX	
PD	26-JUN-2001.
XX	
PF	27-MAR-2000; 2000US-00534639.
XX	
PR	12-NOV-1996; 96US-00747488.
PR	14-OCT-1997; 97US-00950064.
PR	10-FEB-1998; 98US-00021298.
PR	11-MAY-1998; 98US-00075257.
XX	
PA	(DIAM/) DIAMOND D J.
PA	(YORK/) YORK J.
XX	
PI	Diamond DJ, York J;
Pt	
PT	WPI, 2001-431950/46.
XX	
PS	New peptides that are immunogenic epitopes of the human cytomegalovirus
XX	(CMV), useful for activating cytotoxic T lymphocytes (CTL) or CTL
XX	precursors to elicit an immune response against human CMV by normal or
XX	immunodeficient subjects.
XX	
XX	Example 11; Col 16; 17pp; English.
XX	
XX	The invention relates to immunologically active peptides, and functional
XX	variants thereof, capable of eliciting a cellular immune response to
XX	human cytomegalovirus (HCMV) in humans. The peptides are capable of
XX	directing human cytotoxic T lymphocytes (CTL) to recognise and lyse human
XX	cells infected with HCMV. Such immunologically active peptides, in
XX	association with an major histocompatibility complex (MHC) Class I
XX	molecule, are recognised by CTLs of individuals having a latent
XX	(inactive) HCMV infection. Vaccines comprising these peptides are useful
CC	

CC	for activating CTLs and CTL precursors (CTLP), particularly for eliciting
CC	a cellular immune response against HCMV by normal and immunodeficient
CC	subjects. The peptides may be used in an assay for determining the degree
CC	of immunostimulation caused by HCMV. The peptides may also be used for
CC	distinguishing individuals who are seropositive from those who have not
CC	been exposed to HCMV (seronegative individuals). The present sequence is
CC	human cytomegalovirus lipidated vaccine peptide with covalent HTL and CTL
CC	epitopes
XX	
SQ	Sequence 29 AA;
Query Match	59.3%; Score 83; DB 4; Length 29;
Best Local Similarity	69.2%; Pred. No. 4.2e-06;
Matches	18; Conservative 2; Mismatches 6; Indels 0; Gaps 0
OY	3 QYKANSKFITITERGRSLNTVMTL 28 4 QYKANSKFITITEAANLVPMVATV 29
Db	
RESULT 11	
ID	AAE05437 standard; peptide; 29 AA.
XX	
AC	AAE05437;
XX	
DT	24-SEP-2001 (first entry)
XX	
DE	Human cytomegalovirus (HCMV) vaccine peptide #1.
XX	
KM	Human cytomegalovirus; HCMV; infection; vaccine; cytotoxic T lymphocyte;
XX	CTL; CTL precursor; CTLP; Class I antigen;
KW	major histocompatibility complex; MHC; cellular immune response;
XX	CTL activator; CTLP activator; immunostimulant.
OS	Human cytomegalovirus.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1 /note= "Lysine modified with palmitic acid"
FT	Region 4..17 /label= HTL_epitope
FT	Region 18..20 /note= "amino acid spacer"
FT	Region 21..29 /label= CTL_epitope
XX	
PN	US6251399-B1.
XX	
PD	26-JUN-2001.
XX	
PF	27-MAR-2000; 2000US-00534639.
XX	
PR	12-NOV-1996; 96US-00747488.
PR	14-OCT-1997; 97US-00950064.
PR	10-FEB-1998; 98US-00021298.
PR	11-MAY-1998; 98US-00075257.
PA	(DIAM/) DIAMOND D J.
PA	(YORK/) YORK J.
XX	
PI	Diamond DJ, York J;
XX	
DR	WPI; 2001-431950/46.
XX	
PT	New peptides that are immunogenic epitopes of the human cytomegalovirus
PT	(CMV), useful for activating cytotoxic T lymphocytes (CTL) or CTL
PT	precursors to elicit an immune response against human CMV by normal or
PT	immunodeficient subjects.
XX	
PS	Example 11; Col 16; 17pp; English.
CC	The invention relates to immunologically active peptides, and functional

CC variants thereof, capable of eliciting a cellular immune response to
CC human cytomegalovirus (HCMV) in humans. The peptides are capable of
CC directing human cytotoxic T lymphocytes (CTL) to recognise and lyse human
CC cells infected with HCMV. Such immunologically active peptides, in
CC association with an major histocompatibility complex (MHC) Class I
CC molecule, are recognised by CTLs of individuals having a latent
CC (inactive) HCMV infection. Vaccines comprising these peptides are useful
CC for activating CTLs and CTL precursors (CTLp), particularly for eliciting
CC a cellular immune response against HCMV by normal and immunodeficient
CC subjects. The peptides may be used in an assay for determining the degree
CC of immunostimulation caused by HCMV. The peptides may also be used for
CC distinguishing individuals who are seropositive from those who have not
CC been exposed to HCMV (seronegative individuals). The present sequence is
CC human cytomegalovirus lipidated vaccine peptide with covalent HTL and CTL
CC epitopes
SQ Sequence 29 AA;

Query Match 59.3%; Score 83; DB 4; Length 29;
Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QYIKANSKFGITGRSLYNTVATL 28
Db ||||| ||||| : |||||
4 QYIKANSKFGITTEAANLVPMVAVTV 29

RESULT 12

AAU10851
ID AAU10851 standard; peptide; 29 AA.

AC AAU10851;

DT 14-FEB-2002 (first entry)

DE Human cytomegalovirus pp65 protein based lipidated vaccine #3.

XX Immunogenic epitope; pp65; cytotoxic T cell; CTL; HCMV; vaccine;
KM virucide; resteenosis; coronary angioplasty; morbidity; mortality; AIDS;
KM acquired immunodeficiency syndrome; viral retinitis; blindness;
KM encephalitis; enteritis; mononucleosis; interstitial pneumonia;
KM malignancy; Kaposi's sarcoma; post-bone marrow transplant complication.

XX Human cytomegalovirus.
OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Covalently linked to 3 palmitic acid moieties"

FT Peptide 4..17 /label= HTL epitope

FT Peptide /note= "Helper T lymphocyte epitope"

FT Peptide 21..29 /label= CTL epitope

FT Peptide /note= "Cytotoxic T lymphocyte epitope"

PN WO200172782-A2.

PD 04-OCT-2001.

PF 16-MAR-2001; 2001WO-US008576.

PR 27-MAR-2000; 2000US-00534639.

PR 20-OCT-2000; 2000US-00692170.

PA (CITY) CITY OF HOPE.

PI Diamond DJ;

PT New immunogenic peptides useful as a vaccination against cytomegalovirus
infection, comprises epitopes of human cytomegalovirus which are

PT recognized by human cytotoxic T-lymphocytes.
XX Claim 1; Page 26; 65pp; English.

XX The invention relates to immunogenic peptides derived from human
CC cytomegalovirus (HCMV) which are recognised by human cytotoxic T-
CC lymphocytes (CTL). The invention also discloses a vaccine against HCMV
CC comprising one of the peptides, and a pharmaceutically acceptable
CC carrier, a cellular vaccine against HCMV comprising antigen presenting
CC cells that have been treated in vitro so that they express the peptide,
CC and a recombinant viral vector vaccine expressing a gene encoding the
CC peptide. The vaccine is used to modulate the immune response to human
CC cytomegalovirus infection. HCMV is implicated in a number of diseases and
CC conditions including resteenosis after coronary angioplasty,
CC morbidity/mortality in AIDS (acquired immunodeficiency syndrome)
CC patients, viral retinitis leading to blindness, encephalitis, enteritis,
CC mononucleosis, interstitial pneumonia, malignancy (e.g. Kaposi's
CC sarcoma) and post-bone marrow transplant complications. The present
CC sequence is an HCMV pp65 based immunogenic epitope lipidated vaccine of
CC the invention

SQ Sequence 29 AA;

Query Match 59.3%; Score 83; DB 5; Length 29;
Best Local Similarity 69.2%; Pred. No. 4.2e-06;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QYIKANSKFGITGRSLYNTVATL 28
Db ||||| ||||| : |||||
4 QYIKANSKFGITTEAANLVPMVAVTV 29

RESULT 13

AAU10850
ID AAU10850 standard; peptide; 29 AA.

AC AAU10850;

DT 14-FEB-2002 (first entry)

DE Human cytomegalovirus pp65 protein based lipidated vaccine #2.

XX Immunogenic epitope; pp65; cytotoxic T cell; CTL; HCMV; vaccine;
KM virucide; resteenosis; coronary angioplasty; morbidity; mortality; AIDS;
KM acquired immunodeficiency syndrome; viral retinitis; blindness;
KM encephalitis; enteritis; mononucleosis; interstitial pneumonia;
KM malignancy; Kaposi's sarcoma; post-bone marrow transplant complication.

XX Human cytomegalovirus.
OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Covalently linked to 2 palmitic acid moieties"

FT Peptide 4..17 /label= HTL epitope

FT Peptide /note= "Helper T lymphocyte epitope"

FT Peptide 21..29 /label= CTL epitope

FT Peptide /note= "Cytotoxic T lymphocyte epitope"

PN WO200172782-A2.

PD 04-OCT-2001.

PF 16-MAR-2001; 2001WO-US008576.

PR 27-MAR-2000; 2000US-00534639.

PR 20-OCT-2000; 2000US-00692170.

PA (CITY) CITY OF HOPE.

PI Diamond DJ;

XX WPI; 2002-041233/05.
 XX New immunogenic peptides useful as a vaccination against cytomegalovirus
 PT infection, comprises epitopes of human cytomegalovirus which are
 PT recognized by human cytotoxic T-lymphocytes.
 XX
 PS Claim 1; Page 26; 65pp; English.
 XX
 CC The invention relates to immunogenic peptides derived from human
 CC cytomegalovirus (HCMV) which are recognised by human cytotoxic T-
 CC lymphocytes (CTL). The invention also discloses a vaccine against HCMV
 CC comprising one of the peptides, and a pharmaceutically acceptable
 CC carrier, a cellular vaccine against HCMV comprising antigen presenting
 CC cells that have been treated in vitro so that they express the peptide,
 CC and a recombinant viral vector vaccine expressing a gene encoding the
 CC peptide. The vaccine is used to modulate the immune response to human
 CC cytomegalovirus infection. HCMV is implicated in a number of diseases and
 CC conditions including, restenosis after coronary angioplasty,
 CC morbidity/mortality in AIDS (acquired immunodeficiency syndrome)
 CC patients, viral retinitis leading to blindness, encephalitis, enteritis,
 CC mononucleosis, interstitial pneumonia, malignancy (e.g. Kaposi's
 CC sarcoma) and post-bone marrow transplant complications. The present
 CC sequence is an HCMV pp65 based immunogenic epitope lipidated vaccine of
 CC the invention
 XX
 SQ Sequence 29 AA;
 XX
 Query Match 59.3%; Score 83; DB 5; Length 29;
 Best Local Similarity 69.2%; Pred. No. 4.2e-06;
 Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 3 QYKANSKRFITGTERGRSLYNTVAII 28
 |||||
 Db 4 QYKANSKRFITGTEAANLVPWVAIV 29
 |||||
 RESULT 14
 AAU11422
 ID AAU11422 standard; peptide; 28 AA.
 XX
 AC AAU11422;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DB Synthetic immunogen peptide 3.
 XX
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinizing hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX
 OS Clostridium tetani.
 OS Mammalia.
 OS Syntheitic.
 OS Chimeric.
 XX
 XX Key Location/Qualifiers
 FT 1..15 /note= "Tetanus toxoid sequence (830-844 aa)"
 FT Peptide
 FT 16..19 /note= "Spacer peptide"
 FT Peptide
 FT 20..28 /note= "Gonadotropin releasing hormone epitope"
 FT Modified-site 28
 FT /note= "Amlidated glycine or glycineamide"
 XX
 XX WO200185763-A2.
 XX
 XX 15-NOV-2001.
 XX
 XX 04-MAY-2001; 2001WO-US014363.
 XX
 XX

XX 05-MAY-2000; 2000US-0202328P.
 XX
 XX (APHT-) APHTON CORP.
 PA
 XX Grimes S, Michaeli D, Stevens VC;
 PI
 XX WPI; 2002-049440/06.
 XX
 DR Novel synthetic immunogen for inducing immune response against
 XX gonadotropin releasing hormone, comprises fusion peptide having
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.
 XX
 PS Claim 11; Page 8; 43pp; English.
 XX
 CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH also known as
 CC luteinizing hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 SQ Sequence 28 AA;
 XX
 Query Match 58.6%; Score 82; DB 5; Length 28;
 Best Local Similarity 85.0%; Pred. No. 5.9e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 QYKANSKRFITGTERGRSLY 22
 |||||
 Db 1 QYKANSKRFITGTELGPSLH 20
 |||||
 RESULT 15
 ABB79182
 ID ABB79182 standard; peptide; 29 AA.
 XX
 AC ABB79182;
 XX
 DT 07-AUG-2002 (first entry)
 XX
 DB Human cytomegalovirus PPI50 related vaccine peptide SEQ ID NO:8.
 DE
 XX Human cytomegalovirus PPI50; HCMV; human leukocyte antigen; HLA; CTL;
 KW cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
 KW immunogenic; virucide; immune response.
 XX
 OS Human cytomegalovirus.
 OS
 XX WO200234769-A2.
 XX
 XX 02-MAY-2002.
 XX
 XX 22-OCT-2001; 2001WO-US032589.
 XX
 XX 20-OCT-2000; 2000US-0241944P.
 XX
 XX (CITY) CITY OF HOPE.
 PA
 XX Diamond DJ;
 PI
 XX
 XX WPI; 2002-471432/50.
 XX
 XX New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
 PT cytomegalovirus PPI50 useful for creating vaccines against
 PT cytomegalovirus infection.
 XX
 XX

XX Disclosure; Page 14; 28pp; English.

PS The present invention describes a peptide (I) which is an immunogenic
 XX epitope recognised by CD8+ class I major histocompatibility complex (MHC)
 CC restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent
 CC cytomegalovirus (HCMV) infection. Also described are: (1) a vaccine (II)
 CC against human cytomegalovirus (CMV), which comprises (I); (2) a cellular
 CC vaccine (III) against human CMV comprises antigen presenting cells which
 CC present (I); (3) a recombinant viral vector (IV) which contains a gene
 CC encoding (I); and (4) an immunological reagent (V) which comprises (I).
 CC (I) has virucide activity. (II) or (III) are useful for modulating the
 CC immune response to HCMV infection, and for vaccinating a mammal in need
 CC of it against HCMV. (I) is useful for eliciting a cellular immune against
 CC HCMV by normal and immunodeficient subjects, for creating efficient
 CC vaccines against HCMV, and in immunological methods to detect Pp150-
 CC reactive CTL in a patient or a sample from the patient. The present
 CC sequence represents a HCMV Pp150 related vaccine peptide, which is given
 XX in the exemplification of the present invention

SQ Sequence 29 AA;

Query Match

Best Local Similarity 57.9%; Score 81; DB 5; Length 29;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GRQYTKANSKFIGITE 16

DB 14 GRQYTKANSKFIGITE 29

Search completed: July 20, 2004, 06:30:26
 Job time : 43.7705 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 9.63934 Seconds

(Without alignments)
279,413 Million cell updates/sec

Title: US-09-673-166B-274

Perfect score: 144

Sequence: 1 GRQYIKANSKFIGITGRGRIKPEVHGV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	48.6	1315	1 BRLTNT	tentoxilysin (EC 3
2	54	37.5	559	2 B47175	reverse transcript
3	54	37.5	1002	2 S54378	pol polypeptide
4	54	37.5	1003	1 GNWMLV	HIV-1 retropepsin
5	54	37.5	1003	1 B44001	HIV-1 retropepsin
6	54	37.5	1003	2 T09440	pol polypeptide
7	54	37.5	1012	1 GNWMLV	HIV-1 retropepsin
8	54	37.5	1015	1 GNWMLV	HIV-1 retropepsin
9	53	36.8	1002	1 GNWMLV	HIV-1 retropepsin
10	52	36.1	902	2 T01668	HIV-1 retropepsin
11	51	35.4	559	2 A47175	reverse transcript
12	50	34.7	243	2 F96595	unknown protein, 2
13	49	34.0	278	2 G82074	mag2 protein VCC45
14	49	34.0	323	2 B58977	probable ABC trans
15	48	33.3	83	2 B84174	hypothetical prote
16	48	33.3	492	2 E86829	ribosome ABC transp
17	48	33.3	697	1 TVVPTL	large T antigen
18	47	32.6	152	2 A37815	globin (domain E7)
19	47	32.6	352	2 H12894	probable o-sialogl
20	47	32.6	674	2 T20571	hypothetical prote
21	47	32.6	721	2 C84073	DNA topoisomerase
22	46.5	32.3	227	2 AH1667	phosphoribosylform
23	46.5	32.3	287	2 D97164	apases involved i
24	46.5	32.3	564	2 T28949	hypothetical prote
25	46.5	32.3	899	2 T42976	hypothetical prote
26	46	31.9	349	2 T43043	probable acetyl-Co
27	46	31.9	447	2 H97146	sldephore/Surfac
28	46	31.9	448	2 B64503	hypothetical prote
29	46	31.9	493	2 H69689	ribosome ABC transp

30	46	31.9	511	2 E84685	probable seed stor
31	46	31.9	569	2 B82559	30S ribosomal prot
32	46	31.9	728	2 I65413	sodium-dependent n
33	46	31.9	730	2 I52532	sodium-dependent n
34	46	31.9	912	2 S33980	pol polypeptide
35	46	31.9	922	2 AG1827	maltooligosyltreha
36	46	31.9	1003	1 GNWMA2	HIV-1 retropepsin
37	45.5	31.6	455	2 T10406	protein ME53 - Org
38	45.5	31.6	899	2 G36812	hypothetical prote
39	45	31.2	247	2 F47188	MHC class II histo
40	45	31.2	286	2 A95113	telurite resistan
41	45	31.2	288	2 H97981	telurite resistan
42	45	31.2	424	2 G84317	hypothetical prote
43	45	31.2	462	2 G96506	hypothetical prote
44	45	31.2	527	2 T37055	probable oxidoredu
45	45	31.2	689	2 E89898	hypothetical prote

ALIGNMENTS

RESULT 1
BRLTNT
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N:Alternate names: tetanus neurotoxin
C:Species: Clostridium tetani
C:Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.
EMBO J. 5, 2495-2502, 1986
A:Title: Tetanus toxin: Primary structure, expression in E. coli, and homology with botu
A:Reference number: A25689; MUID:87053814; PMID:3536478
A:Accession: A25689
A:Molecule type: DNA
A:Residues: 1-1315 <ERS>
A:Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:940770
R:Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A:Title: The complete nucleotide sequence of tetanus toxin.
A:Reference number: A25757; MUID:87040747; PMID:3774547
A:Accession: A25757
A:Molecule type: DNA
A:Residues: 1-1315 <FAI>
A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774
A:Experimental source: strain CN3911
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in E
A:Reference number: A25194; MUID:86085672; PMID:3510187
A:Accession: A25194
A:Molecule type: DNA
A:Residues: 743-1315 <FA2>
A:Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921
A:Accession: B25194
A:Molecule type: protein
A:Residues: 865-894 <FA3>
R:Matsuda, M.; Iel, D.L.; Sugimoto, N.; Ozutsuni, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A:Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal
A:Reference number: A60759; MUID:90035436; PMID:2478476
A:Accession: A60759
A:Molecule type: protein
A:Residues: 461-475 <MAT>
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A:Reference number: JS0098; MUID:89093918; PMID:2463305
A:Contents: annotation; epitope region
R:Schliavo, G.; Benfenati, F.; Pouliat, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, B.R.
Nature 359, 832-835, 1992
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic
A:Reference number: S27125; MUID:93063293; PMID:1331807
A:Contents: annotation

R:de Filipiis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Bur. J. Biochem. 229, 61-69, 1995
 A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A:Reference number: S69348; MUID:95262688; PMID:7744050
 A:Accession: S69348
 A:Molecule type: protein
 A:Residues: 2-31 <DEF>
 C:Comment: The source of this protein was an extrachromosomal plasmid.
 C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fr
 dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
 C:Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglio
 C:Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
 presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
 C:Function:
 A:Description: blocks neuroexocytosis via hydrolysis of a Glu-Phe peptide bond in synapt
 C:Superfamily: tetanus toxin
 C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F:2-457/Product: tetroxylisin light chain (fragment A) #status predicted <TRL>
 F:461-1315/Product: tetroxylisin heavy chain (fragment B.C) #status experimental <TTH>
 F:461-864/Domaln: channel forming (fragment B) #status predicted <TXB>
 F:865-1315/Domaln: ganglioside binding (fragment C) #status predicted <TXC>
 F:233-237/Binding site: zinc (His) #status predicted
 F:234/Active site: Glu #status predicted

Query Match 48.6%; Score 70; DB 1; Length 1315;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QYKANSKFIIGITE 16
 |||||
 Db 830 QYKANSKFIIGITE 843

RESULT 2

B47175
 Reverse transcriptase, AZT-sensitive variant - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1996
 C:Accession: B47175
 R:Moht, H.; Singh, M.K.; Ching, W.T.; Ho, D.D.
 Proc. Natl. Acad. Sci. U.S.A. 90, 25-29, 1993
 A:Title: Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the
 A:Reference number: A47175; MUID:93126353; PMID:7678340
 A:Accession: B47175
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: DNA
 A:Residues: 1-559 <MOH>
 A:Note: Sequence extracted from NCBI backbone (NCBI:P.122099)
 C:Superfamily: pol polyprotein

Query Match 37.5%; Score 54; DB 2; Length 559;
 Best Local Similarity 45.2%; Pred. No. 6.1;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIIGITERG-----RIIKEPVHGV 28
 |||||
 Db 287 KALTEVIPLEAELELAENREILKEPVHGV 317

RESULT 3

S54378
 pol polyprotein - human immunodeficiency virus type 1 (fragment)
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
 C:Accession: S54378
 R:Theodore, T.; Buckler-White, A.J.
 submitted to the EMBL Data Library, July 1989
 A:Reference number: S54377
 A:Accession: S54378
 A:Status: preliminary
 A:Molecule type: genomic RNA
 A:Residues: 1-1002 <THE>
 A:Cross-references: EMBL:M2639; NID:g329377; PIDN:AAA45366.1; PID:g329381

C:Superfamily: pol polyprotein
 C:Keywords: polyprotein

Query Match 37.5%; Score 54; DB 2; Length 1002;
 Best Local Similarity 45.2%; Pred. No. 11;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIIGITERG-----RIIKEPVHGV 28
 |||||
 Db 441 KALTEVIPLEAELELAENREILKEPVHGV 471

RESULT 4

GNVWIV
 HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate LAV-1a)
 N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Jun-2002
 C:Accession: A03966
 R:Main-Hobson, S.; Santiago, P.; Danos, O.; Cole, S.; Allizon, M.
 Cell 40, 9-17, 1985
 A:Title: Nucleotide sequence of the AIDS virus, LAV.
 A:Reference number: A90865; MUID:8509333; PMID:2981635
 A:Accession: A03966
 A:Molecule type: DNA
 A:Residues: 1-1003 <MAI>
 C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, rev
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polyprotein
 C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot
 F:57-155/Product: retropepsin #status predicted <RTP>
 F:81/Active site: Asp (shared with dimeric partner) #status experimental

Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 11;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIIGITERG-----RIIKEPVHGV 28
 |||||
 Db 442 KALTEVIPLEAELELAENREILKEPVHGV 472

RESULT 5

B44001
 HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (strain YU-2)
 N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
 C:Accession: B44001
 R:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
 J. Virol. 66, 6587-6600, 1992
 A:Title: Complete nucleotide sequence, genome organization, and biological properties of
 A:Reference number: A44001; MUID:93021387; PMID:1404605
 A:Accession: B44001
 A:Molecule type: DNA
 A:Residues: 1-1003 <LIV>
 A:Cross-references: GB:M93258
 C:Comment: This protein is synthesized as a gag-pol polyprotein.
 C:Comment: Specific enzymatic cleavages may yield mature proteins including proteinase, 1
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polyprotein
 C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot
 F:57-155/Product: retropepsin #status predicted <RTP>
 F:81/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 11;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

```

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      A;Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
      A;Reference number: A93353; MUID:85111123; PMID:2578615
      A;Accession: A03965
      A;Molecule type: DNA
      A;Residues: 1-1015 <RAT>
      A;Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA4414
      A;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
      C;Gene: pol
      C;Superfamily: pol polyprotein

RESULT 6
T09440
pol polyprotein - human immunodeficiency virus type 1 (strain JRP1) (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-Jan-2000
C;Accession: T09440
R;Peng, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
Submitted to the EMBL Data Library, July 1996
A;Reference number: Z16673
A;Accession: T09440
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1003 <PAN>
A;Cross-references: EMBL:U63632; NID:91465777; PID:91465779
C;Gene: pol
C;Superfamily: pol polyprotein

Query Match      37.5%; Score 54; DB 2; Length 1003;
Best Local Similarity 45.2%; Pred. No. 11;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      442 KALTEVPLTEAELELAENREILKEPVHGV 472

RESULT 7
GNVWVL
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate IV)
N:Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Jun-2002
C;Accession: A03967
R;Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nucleic Acids Res 13, 450-458, 1985
A;Title: Nucleotide acid structure and expression of the human AIDS/lymphadenopathy retrovi
A;Reference number: A93355; MUID:85111157; PMID:2982104
A;Accession: A03967
A;Molecule type: DNA
A;Residues: 1-1012 <MUE>
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleo
F;6-164/Product: retropepsin #status predicted <RTP>
F;90/Active site: Asp (shared with dimeric partner) #status experimental

Query Match      37.5%; Score 54; DB 1; Length 1012;
Best Local Similarity 45.2%; Pred. No. 12;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      451 KALTEVPLTEAELELAENREILKEPVHGV 481

RESULT 8
GNVWVH3
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate HTIV-III
N:Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 03-Jun-2002
C;Accession: A03965
R;Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora
nberger, D.A.; Papas, T.S.; Ghayab, U.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

```

```

      Nature 313, 277-284, 1985
      A;Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
      A;Reference number: A93353; MUID:85111123; PMID:2578615
      A;Accession: A03965
      A;Molecule type: DNA
      A;Residues: 1-1015 <RAT>
      A;Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA4414
      A;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
      C;Gene: pol
      C;Superfamily: pol polyprotein

Query Match      37.5%; Score 54; DB 1; Length 1015;
Best Local Similarity 45.2%; Pred. No. 12;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      454 KALTEVPLTEAELELAENREILKEPVHGV 484

RESULT 9
GNLJND
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate NDK)
N:Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Jun-2002
C;Accession: J00067
R;Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
Gene 81, 275-284, 1989
A;Title: Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human immuno
A;Reference number: J00065; MUID:90034200; PMID:2806917
A;Accession: J00067
A;Molecule type: DNA
A;Residues: 1-1002 <SPI>
A;Cross-references: GB:M27323; NID:9328154; PIDN:AAA44869.1; PID:9328158
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleo
F;56-154/Product: retropepsin #status predicted <RTP>
F;80/Active site: Asp (shared with dimeric partner) #status predicted

Query Match      36.8%; Score 53; DB 1; Length 1002;
Best Local Similarity 41.9%; Pred. No. 16;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

Qy      6 KANSKFIGITRG-----RIIKEPVHGV 28
      ||::|||
      Db      441 KALTEVPLTEAELELAENREILKEPVHGV 471

RESULT 10
T01668
pol polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T01668
R;Alizon, M.; Main-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A;Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isolat
A;Reference number: Z14389; MUID:86245056; PMID:2424612
A;Accession: T01668
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-902 <ALI>
A;Cross-references: EMBL:K03456; NID:960228; PIDN:CAA28012.1; PID:960230
C;Superfamily: pol polyprotein

```

Query Match 36.1%; Score 52; DB 2; Length 902;
 Best Local Similarity 66.7%; Pred. No. 20;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 14 ITERGRILKEPVHG 28
 : |||||
 Db 357 LAENREILKEPVHG 371

RESULT 11
 A47175
 reverse transcriptase, AZT-resistant variant - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1996
 C:Accession: A47175
 R:Mohtai, H.; Singh, M.K.; Ching, W.T.; Ho, D.D.
 A:Title: Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the
 A:Reference number: A47175; MUID:93126353; PMID:7678340
 A:Accession: A47175
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: nucleic acid
 A:Residues: 1-559 <MOH>
 A:Note: sequence extracted from NCBI backbone (NCBIP:122100)
 C:Superfamily: pol polypeptide

Query Match 35.4%; Score 51; DB 2; Length 559;
 Best Local Similarity 41.9%; Pred. No. 17;
 Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

OY 6 KANSKRIIGTERG-----RIKKEPVHG 28
 |||||
 Db 287 KALTEVIGLTFEAELELAENREILKEPVHG 317

RESULT 12
 P96595
 unknown protein, 25817-24837 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: P96595
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, U.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
 Kerr, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: P96595
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-243 <STO>
 A:Cross-references: GB:AE051173; NID:G10645423; PIDN:AA621540.1, GSPDB:GN00141
 C:Genetics:
 A:Gene: F7A10.13
 A:Map position: 1

Query Match 34.7%; Score 50; DB 2; Length 243;
 Best Local Similarity 50.0%; Pred. No. 9.9;
 Matches 12; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

OY 2 RQYIKANSKRIIGTERGRILKEPV 25
 : |||||
 Db 94 RYTTTRGS-MGIDSGRIKEPV 115

RESULT 13
 G82074

mazG protein VC2450 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 18-Aug-2003
 C:Accession: G82074
 R:Heidelberger, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, U.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.
 I. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: G82074
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-278 <HRI>
 A:Cross-references: GB:AE004315; GB:AE003852; NID:G9657024; PIDN:AA695592.1, GSPDB:GN0012;
 A:Experimental source: serogroup O1, strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC2450
 A:Map position: 1
 C:Superfamily: nucleoside triphosphate pyrophosphohydrolase, MazG type; beta-lactamase re

Query Match 34.0%; Score 49; DB 2; Length 278;
 Best Local Similarity 36.0%; Pred. No. 16;
 Matches 9; Conservative 7; Mismatches 5; Indels 4; Gaps 1;

OY 6 KANSKRI-----GITERGRILKEPVH 26
 |||||
 Db 228 KANAKFRFRFGVEVDKAKAKKPLH 252

RESULT 14
 E95877
 probable ABC transporter periplasmic substrate-binding protein SMO20295 [imported] - Sinc
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 01-Mar-2002
 C:Accession: E95877
 R:Finan, T.M.; Weidner, S.; Wong, K.; Bhurmaster, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95877
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CA048685.1; PID:G15140157; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Aboia, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMO20295
 A:Genome: Plasmid
 C:Superfamily: conserved hypothetical protein H11028

Query Match 34.0%; Score 49; DB 2; Length 323;
 Best Local Similarity 40.7%; Pred. No. 19;
 Matches 11; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 1 GROVTKANSKRIIGTERGRILKEPVHG 27
 : |||||
 Db 30 GMOGTAGDPQPEGVTEARITIKESGG 56

RESULT 15
 B84174
 hypothetical protein Vng0138h [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: B84174
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A;Title: Genome sequence of Halobacterium species NRC-1.
 A;Reference number: A84160; MUID:20504483; PMID:11016950
 A;Accession: B84174
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-83 <STO>
 A;Cross-references: GB:AE004437; NID:G10579784; PIDN:AA618758.1; GSPDB:GN00138
 C;Genetics:
 A;Gene: VMG0138H

Query Match 33.3%; Score 48; DB 2; Length 83;
 Best Local Similarity 50.0%; Pred. No. 6.2;
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RQYXANSKFTGTERGR 19
 | : : : : : | | | | |
 Db 35 RRAVEASAKVTGAERGR 52

Search completed: July 20, 2004, 06:34:10
 Job time : 11.6393 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 5.5082 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166B-274

Sequence: 144
1 GROYIKANSKFGITGRGRIKLEPHGV 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	70	46.6	1314	1	TEXC_CLOTE
2	54	37.5	1002	1	P04589 human immun
3	54	37.5	1002	1	P04589 human immun
4	54	37.5	1003	1	P04585 human immun
5	54	37.5	1003	1	P04585 human immun
6	54	37.5	1003	1	P04585 human immun
7	54	37.5	1006	1	P04585 human immun
8	54	37.5	1015	1	P04585 human immun
9	54	37.5	1015	1	P04585 human immun
10	54	37.5	1015	1	P04585 human immun
11	54	37.5	1015	1	P04585 human immun
12	53	36.8	1002	1	P04585 human immun
13	53	36.8	1002	1	P04585 human immun
14	52	36.1	1002	1	P04585 human immun
15	52	36.1	1002	1	P04585 human immun
16	52	36.1	1007	1	P04585 human immun
17	49	34.0	1007	1	P04585 human immun
18	49	34.0	1002	1	P04585 human immun
19	48	33.3	697	1	TALA_POVLY
20	47	32.6	152	1	GLB7_ARTSX
21	47	32.6	352	1	GCP_TREPA
22	46.5	32.3	227	1	PURQ_LISTN
23	46	31.9	448	1	Y628_METIA
24	46	31.9	493	1	RBSA_BACVN
25	46	31.9	729	1	NTT7_BOVIN
26	46	31.9	729	1	NTT7_BOVIN
27	46	31.9	730	1	NTT7_BOVIN
28	46	31.9	1003	1	POL_HV1A2
29	45.5	31.6	455	1	ME53_NPVOP
30	45.5	31.6	899	1	ME53_NPVOP
31	45	31.2	207	1	V120_HSVSA
32	45	31.2	384	1	PEH1_BUCBP
33	44.5	30.9	288	1	ALF_MYCGE

34	44.5	30.9	585	1	SP2E_BACME	P49600 bacillus me
35	44.5	30.9	1682	1	MSPI_PLAFC	P19588 plasmodium
36	44.5	30.9	1701	1	MSPI_PLAFC	P13819 plasmodium
37	44.5	30.9	1701	1	MSPI_PLAFC	P08569 plasmodium
38	44.5	30.9	1726	1	MSPI_PLAFC	P04934 plasmodium
39	44.5	30.9	1726	1	MSPI_PLAFC	P04935 plasmodium
40	44	30.6	66	1	VG84_BPMLS	005301 mycobacteri
41	44	30.6	333	1	ILVC_METHH	027491 methanobact
42	44	30.6	354	1	RECA_ACTFA	P34716 acidiphilin
43	44	30.6	364	1	ARGD_PYRAB	09phk2 xylella fas
44	44	30.6	660	1	TPLE_XLFA	P43939 haemophilus
45	43.5	30.2	106	1	Y094_HAEIN	

ALIGNMENTS

RESULT 1
ID TEXC_CLOTE STANDARD; PRT; 1314 AA.
AC P04589;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:
DE tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TEXC OR CTP60.
OS Clostridium tetani.
OG Plasmid p88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudek M., Habermann E., Niemann H.,
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.,
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=PE88;
RX MEDLINE=22457253; PubMed=12552129;
RA Breigemann H., Baewer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.,
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RN [4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.,
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegstein K., Henschen A., Weller U., Habermann E.,
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";

RN Eur. J. Biochem. 188:39-45 (1990) .
 RN [6]
 RN PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henschen A.H., Weller U., Hebertmann E.;
 RT "limited proteolysis of tetanus toxin. Relation to activity and
 identification of cleavage sites.";
 RL Eur. J. Biochem. 202:41-51 (1991) .
 RN [7]
 RN IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "Tetanus toxin is a zinc protein and its inhibition of
 neurotransmitter release and protease activity depend on zinc.";
 RL EMO J. 11:3577-3583 (1992) .
 RN [8]
 RN IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=91063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.F.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835 (1992) .
 RN [9]
 RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792 (1997) .
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 BOND OF SYNAPTOBREVIN-2.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
 synaptobrevin 2.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 YIELD SUBUNITS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 AND ARE NON-TOXIC AFTER SEPARATION.
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 GANGLIOSIDE RECEPTORS.
 CC -1- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X04436; CAA28033.1; -
 DR EMBL: X06214; CAA29564.1; -
 DR EMBL: AF528097; AA037454.1; -
 DR EMBL: M12739; AAA23282.1; -
 DR PIR: A25689; BTCLTN.
 DR PDB: 1AF9; 29-APR-98.
 DR PDB: 1ABD; 14-OCT-98.
 DR PDB: 1DOH; 27-MAR-00.
 DR PDB: 1DFQ; 24-MAR-00.
 DR PDB: 1DIW; 24-MAR-00.
 DR PDB: 1DLI; 24-MAR-00.
 DR PDB: 1FV3; 05-SEP-01.
 DR MEROPS: M27.001; -
 DR InterPro: IPR008985; Consil like Jec_g1.
 DR InterPro: IPR002160; Kunitz_legume.
 DR InterPro: IPR006025; Pept_M_Zn_BS

[illegible]

Query Match 48.6%; Score 70; DB 1; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 QYIKANSKFIGITE 16
 |||||
 Db 829 QYIKANSKFIGITE 842

RESULT 2
 POL_HV1EL STANDARD; PRT; 1002 AA.
 ID POL_HV1EL
 AC P04589; 077906;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (H1 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86245056; PubMed=2424612;
 RA Alison M., Main-Hobson S., Montagnier L., Sonigo P.;
 RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
 of two isolates from African patients."
 RL Cell 46:63-74(1986).
 CC -1 CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 and P1' variable, but often Pro.
 CC -1 CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester.
 CC -1 CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + {DNA} (N).
 CC -1 PTM: Cleavage sites that yield the mature proteins remain to be
 determined.
 CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 KNOWN AS THE RETROPEPSIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; A07108; CA00612.1; -;
 CC HSMB; K03454; AAA44325.1; ALT_INIT.
 CC HSSP; P03366; IHMV.
 CC HIV; K03454; POLSELI.
 CC MEROPS; A02.001; -;
 CC InterPro: IPR001969; Aspprotease_AS.
 CC InterPro: IPR001037; Integrase_C.
 CC InterPro: IPR003308; Integrase_Zn.
 CC InterPro: IPR009007; Pept_A_acid.
 CC InterPro: IPR001995; Peptidase_A2.
 CC InterPro: IPR002156; RNaseH.
 CC InterPro: IPR001584; Rye.
 CC InterPro: IPR000477; RVTse.
 CC Pfam; PF00552; Integrase; 1.
 CC Pfam; PF02022; Integrase_Zn; 1.
 CC Pfam; PF00075; RNaseH; 1.
 CC Pfam; PF00665; rve; 1.
 CC Pfam; PF00077; rvp; 1.
 CC Pfam; PF00078; rvt; 1.
 CC PROSITE; PS00141; ASP_PROTEASE; 1.
 CC PROSITE; PS50175; ASP_PROT_RETROV; 1.
 CC K1M; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
 CC Nuclease; Transferase; RNA-directed DNA polymerase.
 CC PT CHAIN 56 154 PROTEASE.

FT ACT SITE 80 80 BY SIMILARITY.
 SQ SEQUENCE 1002 AA; 114002 MW; 5700903B6893B54 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1002;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGITERG-----RIIKEPVHGV 28
 ||::|||
 Db 441 KALTEVIFLTERAELELAENRILKEPVHGV 471

RESULT 3
 POL_HV1Z2 STANDARD; PRT; 1002 AA.
 ID POL_HV1Z2
 AC P12499;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (Z2/CDC-Z34 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11683;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Theodore T., Buckler-White A.;
 RT Submitted (NOV-1989) to the HIV data bank.
 CC -1 CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 and P1' variable, but often Pro.
 CC -1 CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester.
 CC -1 CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + {DNA} (N).
 CC -1 PTM: Cleavage sites that yield the mature proteins remain to be
 determined.
 CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 KNOWN AS THE RETROPEPSIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M22639; AAA45366.1; -;
 CC PIR; S54378; S54378.
 CC PDB; 1E28; 12-SEP-00.
 CC PDB; 1HXW; 04-FEB-98.
 CC PDB; 1PRO; 17-AUG-96.
 CC PDB; 1V1J; 13-JAN-99.
 CC PDB; 1V1K; 13-JAN-99.
 CC HIV; M22639; POL\$Z2Z6.
 CC MEROPS; A02.001; -;
 CC InterPro: IPR001969; Aspprotease_AS.
 CC InterPro: IPR001037; Integrase_C.
 CC InterPro: IPR003308; Integrase_Zn.
 CC InterPro: IPR009007; Pept_A_acid.
 CC InterPro: IPR001995; Peptidase_A2.
 CC InterPro: IPR002156; RNaseH.
 CC InterPro: IPR001584; Rye.
 CC InterPro: IPR000477; RVTse.
 CC Pfam; PF00552; Integrase; 1.
 CC Pfam; PF02022; Integrase_Zn; 1.
 CC Pfam; PF00075; RNaseH; 1.
 CC Pfam; PF00665; rve; 1.
 CC Pfam; PF00077; rvp; 1.
 CC Pfam; PF00078; rvt; 1.
 CC PROSITE; PS00141; ASP_PROTEASE; 1.
 CC PROSITE; PS50175; ASP_PROT_RETROV; 1.
 CC DR

KW AIDS; Polypeptide; Hydrolase; Aspartyl protease; Endonuclease;
 KM Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT CHAIN 56 154
 FT ACT SIRE 80 80 BY SIMILARITY.
 SQ SEQUENCE 1002 AA; 113724 MW; CB4AAC9AB4742315 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1002;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKRGIGTERG-----RIIKRPYGV 28
 Db 441 KALTEVIFPLEBAHEALFNENRRIKPEVHGV 471
 RESULT 4
 ID POL_HV1H2 STANDARD; PRT; 1003 AA.
 AC P04585; Q09777; Q9W0C5;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL polypeptide [contains: Protease (Retropesin) (EC 3.4.23.16);
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxId=11706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87299196; PubMed=3040055;
 RA Rather L.C., Fisher A., Jagodzinski L.L., Mitsuya H., Liu R.-S.,
 RA Gallo R.C., Wong-Staal F.;
 RT "Complete nucleotide sequences of functional clones of the AIDS
 RT virus.";
 RT AIDS Res. Hum. Retroviruses 3:57-69(1987).
 RN [2]
 RP REVISIONS.
 RA Ogata N., Alter H.J., Miller R.H., Purcell R.H.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chappey C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
 RX MEDLINE=99043699; PubMed=827997;
 RA Keränen J., Lubkowski J., Zdanov A., Bhatt D., Dunn B.M., Hu K.Y.,
 RA Powell D.J., Kay J., Wlodawer A., Gutschina A.;
 RT "Toward a universal inhibitor of retroviral proteases: comparative
 RT analysis of the interactions of LP-130 complexed with proteases from
 RT HIV-1, FIV, and EIAV.";
 RL Protein Sci. 7:2314-2323(1998).
 RN [5]
 RP STRUCTURE BY NMR OF 57-155.
 RX MEDLINE=97022126; PubMed=866486;
 RA Yamazaki T., Hinck A.P., Wang Y.-X., Nicholson L.K., Torchia D.A.,
 RA Wingfield P., Stahl S.J., Kaufman J.D., Chang C.-H., Donatelli P.U.,
 RA Lam P.Y.S.;
 RT "Three-dimensional solution structure of the HIV-1 protease complexed
 RT with DM9323, a novel cyclic urea-type inhibitor, determined by
 RT nuclear magnetic resonance spectroscopy.";
 RL Protein Sci. 5:495-506(1996).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 156-595.
 RX MEDLINE=96097398; PubMed=8535785;
 RA Ren J., Esnouf R.M., Hopkins A.L., Ross C., Jones E.Y., Stammers D.K.,
 RA Stuart D.I.;
 RT "The structure of HIV-1 reverse transcriptase complexed with
 RT 9-chloro-TIBO: lessons for inhibitor design.";
 RL Structure 3:915-926(1995).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS) OF 156-595.

RX MEDLINE=96208551; PubMed=8648598;
 RA Hopkins A.L., Ren J., Esnouf R.M., Willcox B.B., Jones E.Y., Ross C.,
 RA Miyasaka T., Walker R.T., Tanaka H., Stammers D.K., Stuart D.I.;
 RT "Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT
 RT series reveal conformational changes relevant to the design of potent
 RT non-nucleoside inhibitors.";
 RL J. Med. Chem. 39:1589-1600(1996).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 157-694.
 RX MEDLINE=97268683; PubMed=9108091;
 RA Esnouf R.M., Ren J., Hopkins A.L., Ross C.K., Jones E.Y.,
 RA Stammers D.K., Stuart D.I.;
 RT "Unique features in the structure of the complex between HIV-1
 RT reverse transcriptase and the bis(heteroaryl)piperazine (BHP)
 RT U-90152 explain resistance mutations for this nonnucleoside
 RT inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:3984-3989(1997).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 157-694.
 RX MEDLINE=98356189; PubMed=9689112;
 RA Ren J., Esnouf R.M., Hopkins A.L., Jones E.Y., Kirby I., Keeling J.,
 RA Ross C.K., Larder B.A., Stuart D.I., Stammers D.K.;
 RT "3'-Azido-3'-deoxythymidine drug resistance mutations in HIV-1
 RT reverse transcriptase can induce long range conformational changes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:9518-9523(1998).
 CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
 CC and PI variable, but often pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA) (N).
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K03455; AAB50259.1; ALT_INIT.
 DR EMBL; AF033819; AAC82598.2; -.
 DR PDB; 1REV; 14-OCT-96.
 DR PDB; 1RT1; 21-APR-97.
 DR PDB; 1RT2; 21-APR-97.
 DR PDB; 1RT3; 16-FEB-99.
 DR PDB; 1RT4; 03-APR-96.
 DR PDB; 1RT1; 03-APR-96.
 DR PDB; 1RT0; 03-APR-96.
 DR PDB; 1BVE; 17-AUG-96.
 DR PDB; 1BVG; 17-AUG-96.
 DR PDB; 1KLM; 18-MAR-98.
 DR PDB; 1A30; 29-APR-98.
 DR PDB; 1ODY; 16-FEB-99.
 DR PDB; 1BV7; 14-JAN-00.
 DR PDB; 1BV9; 12-JAN-00.
 DR PDB; 1BNA; 12-JAN-00.
 DR PDB; 1BMB; 12-JAN-00.
 DR PDB; 1COT; 19-JUL-00.
 DR PDB; 1COT; 19-JUL-00.
 DR PDB; 1COT; 19-JUL-00.
 DR PDB; 1C1B; 21-JUL-00.
 DR PDB; 1C1C; 06-SEP-00.
 DR PDB; 1DMF; 12-NOV-97.
 DR PDB; 1DTP; 02-MAR-00.
 DR PDB; 1EP4; 27-SEP-00.
 DR PDB; 1EP4; 27-SEP-00.
 DR PDB; 1EX4; 26-JUL-00.
 DR PDB; 1EXQ; 03-NOV-00.
 DR PDB; 1FK9; 28-FEB-01.

DR PDB; 1FKO; 28-FEB-01.
 DR PDB; 1FKP; 03-NOV-00.
 DR PDB; 1HVA; 06-APR-99.
 DR PDB; 1HVR; 15-MAY-95.
 DR PDB; 1HWR; 23-MAR-99.
 DR PDB; 1HXB; 12-MAR-97.
 DR PDB; 1JXH; 03-OCT-01.
 DR PDB; 1JLA; 03-OCT-01.
 DR PDB; 1JLB; 03-OCT-01.
 DR PDB; 1JLC; 03-OCT-01.
 DR PDB; 1JLF; 03-OCT-01.
 DR PDB; 1JLG; 03-OCT-01.
 DR PDB; 1JLQ; 22-AUG-01.
 DR PDB; 1LM0; 30-OCT-02.
 DR PDB; 1LM2; 30-OCT-02.
 DR PDB; 1LMC; 30-OCT-02.
 DR PDB; 1LWE; 30-OCT-02.
 DR PDB; 1LWF; 30-OCT-02.
 DR PDB; 1MER; 15-APR-98.
 DR PDB; 1MES; 15-APR-98.
 DR PDB; 1MEU; 15-APR-98.
 DR PDB; 1OIW; 18-FEB-03.
 DR PDB; 1ODM; 01-APR-97.
 DR PDB; 1OBS; 15-OCT-97.
 DR PDB; 1OBR; 15-OCT-97.
 DR PDB; 1OBT; 15-OCT-97.
 DR PDB; 1OBU; 15-OCT-97.
 DR PDB; 1RT4; 29-JUL-99.
 DR PDB; 1RT5; 29-JUL-99.
 DR PDB; 1RT7; 29-JUL-99.
 DR PDB; 1RTD; 12-JAN-00.
 DR PDB; 1VAT; 03-APR-96.
 DR PDB; 3PHV; 15-JAN-92.
 DR HIV; K03455; POLSHXB2.
 DR MEROPS; A02.001; -.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00552; Integrase_1.
 DR Pfam; PF02022; Integrase_Zn; 1.
 DR Pfam; PF00075; rnaaseh; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvt; 1.
 DR PROSITE; PS00141; ASP_PROTASE; 1.
 DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
 DR AIBS; Polypeptide; Hydrolase; Aspartyl protease; Endonuclease;
 KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT CHAIN 57 155
 FT ACT_SITE 81 81
 FT STRAND 58 59
 FT STRAND 66 71
 FT TURN 72 73
 FT STRAND 74 80
 FT TURN 82 83
 FT STRAND 88 89
 FT STRAND 99 104
 FT STRAND 109 122
 FT STRAND 126 133
 FT STRAND 140 141
 FT STRAND 143 146
 FT HELIX 147 150
 FT TURN 152 154
 FT STRAND 167 167
 Query Match 37.5%; Score 54; DB 1; Length 1003;

Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 Qy 6 KANSKFIGITERG-----RLKEPVHGV 28
 Db 442 KALTEVIVPLEAELELANREILKEPVHGV 472
 RESULT 5
 POL_HVIOY
 ID POL_HVIOY STANDARD; PRT; 1003 AA.
 AC P20892;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE POL polypeptide [contains: Protease (Retropepsin) (EC 3.4.23.16);
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90148544; PubMed=2559749;
 RA Hue T.; Dazza M.C.; Brun-Vezinet F.; Roelants G.F.; Wain-Hobson S.;
 RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
 RT individual presenting an atypical western blot."
 RL AIDS 3:707-715(1989).
 CC -1- CATALYTIC ACTIVITY: Specific for a PL residue that is hydrophobic,
 CC and PL variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA}(N).
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
 CC HEALTHY GABONESE INDIVIDUAL.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M26727; AAA83392.1; -.
 CC HSSP; P03366; IRVL.
 CC HIV; M26727; POLSOYI.
 DR MEROPS; A02.001; -.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00552; Integrase_1.
 DR Pfam; PF02022; Integrase_Zn; 1.
 DR Pfam; PF00075; rnaaseh; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvt; 1.
 DR PROSITE; PS00141; ASP_PROTASE; 1.
 DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
 DR AIBS; Polypeptide; Hydrolase; Aspartyl protease; Endonuclease;
 KW Nuclease; Transferase; RNA-directed DNA polymerase.
 FT CHAIN 57 155
 FT ACT_SITE 81 81
 FT BY SIMILARITY.

SQ SEQUENCE 1003 AA; 113718 MW; AFB997A0EDB88A98 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGTTERG-----RIIKEPVHGV 28
 DB 442 KALTEVITPLEBAEELAELENREILKEPVHGV 472
 RESULT 6
 POL_HV1Y2 STANDARD; PRT; 1003 AA.
 AC P35963;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93021387; PubMed=1404605;
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
 Shaw G.M.;
 RT "Complete nucleotide sequence, genome organization, and biological
 properties of human immunodeficiency virus type 1 in vivo: evidence
 for limited defectiveness and complementation.";
 RL J. Virol. 66:6587-6600(1992).
 CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 determined.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M93258; -; NOT_ANNOTATED_CDS.
 DR PIR; B44001; B44001.
 DR PDB; 1K6C; 06-FEB-02.
 DR PDB; 1K6F; 06-FEB-02.
 DR PDB; 1K6T; 06-FEB-02.
 DR PDB; 1K6V; 06-FEB-02.
 DR PDB; 1WJA; 13-MAY-98.
 DR PDB; 1WJC; 13-MAY-98.
 DR MEROPS; A02.001; -;
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rye.
 DR InterPro; IPR000477; RYase.
 DR Pfam; PF00552; Integrase; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.

DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rve; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 KW Aids; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
 KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 FT CHAIN 57 155
 FT ACT SITE 81 81
 FT ACT SITE 81 155
 SQ SEQUENCE 1003 AA; 113794 MW; D2F7187FE4E8F49 CRC64;
 Query Match 37.5%; Score 54; DB 1; Length 1003;
 Best Local Similarity 45.2%; Pred. No. 3.3;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGTTERG-----RIIKEPVHGV 28
 DB 442 KALTEVITPLEBAEELAELENREILKEPVHGV 472
 RESULT 7
 POL_HV1MN STANDARD; PRT; 1006 AA.
 AC P05961;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11696;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88219542; PubMed=3369091;
 RA Garago C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
 RA Farrell K., Wong-Staal F., Gallo R.C., Reltz M.S., Jr.;
 RT "Envelope sequences of two new United States HIV-1 isolates.";
 RL Virology 164:531-536(1988).
 CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
 and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 determined.
 CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
 PATIENT IN 1984.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M17449; -; NOT_ANNOTATED_CDS.
 DR HSP; P03366; IRVN.
 DR HIV; M17449; POLSMN.
 DR MEROPS; A02.001; -;
 DR InterPro; IPR001969; Asparticase_AS.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_Zn.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001995; Peptidase_A2.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR001584; Rye.
 DR InterPro; IPR000477; RYase.
 DR Pfam; PF00552; Integrase; 1.

RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 594-729.
 RX MEDLINE=91188281; PubMed=11707186;
 RA Davies J.F. II, Hostomska Z., Hostomsky Z., Jordan S.R.,
 RA Matthews D.A.;
 RT "Crystal structure of the ribonuclease H domain of HIV-1 reverse
 RT transcriptase.";
 RL Science 252:88-95(1991).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 168-723.
 RX MEDLINE=93317673; PubMed=7687065;
 RA Jacobo-Molina A., Ding Y., Nami R.G., Clark A.D. Jr., Lu X.,
 RA Tantillo C., Williams R.L., Kamer G., Ferris A.L., Clark P., Hizi A.,
 RA Hughes S.H., Arnold E.;
 RT "Crystal structure of human immunodeficiency virus type 1 reverse
 RT transcriptase complexed with double-stranded DNA at 3.0-A resolution
 RT shows bent DNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6320-6324(1993).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 168-723.
 RX MEDLINE=92311654; PubMed=1377403;
 RA Kohlstaedt L.A., Wang J., Friedman J.M., Rice P.A., Steitz T.A.;
 RT "Crystal structure at 3.5-A resolution of HIV-1 reverse transcriptase
 RT complexed with an inhibitor.";
 RL Science 256:1783-1790(1992).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF 168-727.
 RX MEDLINE=95166801; PubMed=7532306;
 RA Rodgers D.W., Gambelin S.J., Harris B.A., Ray S., Culp J.S.,
 RA Helling B., Wolf D.J., Debouck C., Harrison S.C.;
 RT "The structure of unliganded reverse transcriptase from the human
 RT immunodeficiency virus type 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:1222-1226(1995).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 168-725.
 RX MEDLINE=95338599; PubMed=7542140;
 RA Ding Y., Das K., Tantillo C., Zhang W., Clark A.D. Jr., Jessen S.,
 RA Lu X., Hsiao Y., Jacobo-Molina A., Andries K., Et A.L.;
 RT "Structure of HIV-1 reverse transcriptase in a complex with the non-
 RT nucleoside inhibitor alpha-APA R 95845 at 2.8-A resolution.";
 RL Structure 3:365-379(1995).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 168-723.
 RX MEDLINE=9643430; PubMed=8805568;
 RA Hsiao Y., Ding Y., Das K., Clark A.D. Jr., Hughes S.H., Arnold E.;
 RT "Structure of unliganded HIV-1 reverse transcriptase at 2.7-A
 RT resolution: implications of conformational changes for polymerization
 RL and inhibition mechanisms.";
 RL Structure 4:853-860(1996).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 168-723.
 RX MEDLINE=9903049; PubMed=9813120;
 RA Hsiao Y., Das K., Ding J., Clark A.D. Jr., Klein J.P., Rosner M.,
 RA Winkler I., Ries G., Hughes S.H., Arnold E.;
 RT "Structure of Tyr188Leu mutant and wild-type HIV-1 reverse
 RT transcriptase complexed with the non-nucleoside inhibitor HBY 097:
 RT inhibitor flexibility is a useful design feature for reducing drug
 RT resistance.";
 RL J. Mol. Biol. 284:313-323(1999).
 RN [12]
 RP CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
 RP and P1' variable, but often Pro.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA] (N).
 CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2, ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its


```

DR PDB; 1HVE; 15-JAN-95.
DR PDB; 1HVS; 14-FEB-95.
DR PDB; 1KIT; 10-JUL-02.
DR PDB; 1KIU; 10-JUL-02.
DR PDB; 1K2B; 10-JUL-02.
DR PDB; 1K2C; 10-JUL-02.
DR PDB; 1ODX; 01-APR-97.
DR PDB; 1TCX; 07-DEC-96.
DR PDB; 1WDE; 16-DEC-98.
DR PDB; 1WDE; 16-DEC-98.
DR PDB; 2BPV; 23-FEB-99.
DR PDB; 2BPW; 23-FEB-99.
DR PDB; 2BPX; 23-FEB-99.
DR PDB; 2BPZ; 23-FEB-99.
DR HIV; K02012; POLSBSH5.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR AIDS; Polypotein; Hydrolyase; Aspartyl protease; Endonuclease;
DR Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93
SQ SEQUENCE 1015 AA; 115015 MM; F79FCB4A1A92CEE CRC64;

Query Match 37.5%; Score 54; DB 1; Length 1015;
Best Local Similarity 45.2%; Pred. No. 3.4;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKITGITEKG-----RLKEPVHGV 28
Db 454 KALTEVIPLEBAELELAENREILKEPVHGV 484

RESULT 10
POL_HY1BR STANDARD; PRT; 1015 AA.
AC P03367;
AT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL. polyprotein [contains: Processase (Retropesin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11686;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=8509333; PubMed=2981635;
RT "Wall-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.",
RL Cell 40:9-17(1985).
RN [2]
RP REVISIONS TO 23-35.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients.",

```

```

RL Cell 46:63-74(1986).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=92190341; PubMed=1799632;
RA Spinellet S., Liu Q.Z., Alzari P.M., Harel P.H., Poljak R.V.;
RT "The three-dimensional structure of the aspartyl protease from the
RT HIV-1 isolate BRU.";
RL Biochimie 73:1391-1396(1991).
CC -1 CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often pro.
CC -1 CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1 CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1 PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02013; -; NOT_ANNOTATED_CDS.
DR PDB; 1HHP; 15-OCT-92.
DR PDB; 1ABG; 15-JUL-98.
DR PDB; 1A94; 16-FEB-99.
DR PDB; 1D4S; 12-OCT-99.
DR PDB; 1D4Y; 13-OCT-99.
DR PDB; 1DAZ; 03-MAY-00.
DR PDB; 1DIF; 08-MAR-96.
DR PDB; 1HOS; 31-OCT-93.
DR PDB; 1HPO; 21-APR-97.
DR PDB; 1HPS; 15-JAN-95.
DR PDB; 1HPX; 08-MAR-96.
DR PDB; 1HSG; 03-APR-96.
DR PDB; 1HRE; 31-JUL-94.
DR PDB; 1HVL; 30-APR-94.
DR PDB; 1UBJ; 14-OCT-96.
DR PDB; 2UBJ; 14-OCT-96.
DR PDB; 7UBJ; 21-APR-97.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR AIDS; Polypotein; Hydrolyase; Aspartyl protease; Endonuclease;
DR Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93
FT STRAND 78 82
FT STRAND 87 92
FT TURN 94 95
FT STRAND 100 102
FT STRAND 111 117
FT TURN 118 119
FT STRAND 120 134

```

```

FT TURN 135 136
FT STRAND 137 146
FT STRAND 152 153
FT HELIX 155 161
FT TURN 162 162
SQ SEQUENCE 1015 AA, 115031 MW, 164702F074A84394 CRC64;

Query Match 37.5%; Score 54; DB 1; Length 1015;
Best Local Similarity 45.2%; Pred. No. 3.4;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIKKEPVGV 28
DB 454 KALTEVIPLEAEALAEAREILKEPVGV 484

RESULT 11
POL HV1PV STANDARD; PRT; 1015 AA.
AC P03368;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI Taxid=11700;
RN RP SEQUENCE FROM N.A.
RX MEDLINE=65111157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabralilla C.D., Benton C.V., Laeky L.A.,
RA Capon D.J.;
RT AIDS/Lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).

[2]
REVISION.
RA Muesing M.A.;
RL Submitted (XXX-1987) to the HIV data bank.
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonooxester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02083; AAB59867.1; -.
CC EMBL; X01762; -. NOT ANNOTATED_CDS.
CC PIR; A03967; GNTWVL.
CC PDB; 1A9M; 17-JUN-98.
CC PDB; 1A9M; 20-AUG-97.
CC PDB; 1A9M; 17-SEP-97.
CC PDB; 1G35; 06-JUN-01.
CC PDB; 1G35; 08-NOV-96.
CC PDB; 1HTG; 31-JUL-94.
CC PDB; 1HTG; 30-APR-94.
CC PDB; 1KJH; 06-MAR-94.
CC PDB; 1NPV; 04-FEB-03.
CC PDB; 1NPV; 04-FEB-03.

```

```

DR HIV; K02083; POL$PV22.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RNase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF002022; Integrase; 1.
DR Pfam; PF000075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rve; 1.
DR Pfam; PF00078; rve; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR AIDS; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
DR Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93 BY SIMILARITY.
SQ SEQUENCE 1015 AA; 115090 MW; 51529D18AC2AF89 CRC64;

Query Match 37.5%; Score 54; DB 1; Length 1015;
Best Local Similarity 45.2%; Pred. No. 3.4;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIKKEPVGV 28
DB 454 KALTEVIPLEAEALAEAREILKEPVGV 484

RESULT 12
POL HV1ND STANDARD; PRT; 1002 AA.
AC P18802;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI Taxid=11695;
RN RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus.";
RL Gene 81:275-284(1989).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonooxester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M27323; AAA44869.1; -.
CC PIR: J00067; GNJND.
CC HSP: P03366; IHMV.
CC HIV: M27323; POL$NDK.
CC MEROPS: A02.001; -.
CC InterPro: IPR001969; Asprotease_AS.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR00308; Integrase_Zn.
CC InterPro: IPR009007; Pept_A_acid.
CC InterPro: IPR001995; Peptidase_A2.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR001584; Rve.
CC InterPro: IPR000477; RVse.
CC Pfam: PF00552; Integrase_1.
CC Pfam: PF02022; Integrase_Zn; 1.
CC Pfam: PF00075; rnaseH; 1.
CC Pfam: PF00665; rve; 1.
CC Pfam: PF00077; rvt; 1.
CC Pfam: PF00078; rvt; 1.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
CC PROSITE: PS50175; ASP_PROT_RETROV; 1.
CC AIDS: Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC CHAIN 56 154
CC ACT SITE 80 80 BY SIMILARITY.
CC SEQUENCE 1002 AA; 113621 MW; 58D868C9896CC1 CRC64;

Query Match 36.8%; Score 53; DB 1; Length 1002;
Best Local Similarity 41.9%; Pred. No. 4.7;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITERG-----RIKKEPVHG 28
   ||:::|||||
Db 441 KALTEVPLTBEAELELAENREILKEPVHG 471

RESULT 13
POL_HY1N5 STANDARD; PRT; 1003 AA.
ID POL_HY1N5
AC P12497;
DT 01-OCT-1989 (Rel. 12; Created)
DT 01-OCT-1989 (Rel. 12; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE POL polypeptide (Contains: Protease (Retropesin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11698;
RN [1]
RP SEQUENCE FROM N.A. (CLONE PML4-3).
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
RL Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.
RV [2]
RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
RA MEDLINE=90354401; PubMed=2201682;
RA Fitzgerald P.M.D., McKeever B.M., van Middlesworth J.F.,
RA Springer J.P., Heimback J.C., Liu C.-T., Herber W.K., Dixon R.A.F.,
RA Darke P.L.;
RA "Crystallographic analysis of a complex between human
RA immunodeficiency virus type 1 protease and acetyl-peptatin at 2.0-A
RA resolution.";
RL J. Biol. Chem. 265:14209-14219(1990).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester..
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
```

```
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M19921; AAA4968.1; -.
CC PDB: 5HP; 15-OCT-91.
CC PDB: 4PHV; 31-OCT-93.
CC PDB: 1B9F; 19-JUL-99.
CC PDB: 1BHL; 28-OCT-98.
CC PDB: 1B14; 18-NOV-98.
CC PDB: 1B1S; 16-SEP-98.
CC PDB: 1B1U; 16-SEP-98.
CC PDB: 1WJB; 13-MAY-98.
CC PDB: 1WJD; 13-MAY-98.
CC PDB: 2ITG; 12-MAR-97.
CC PDB: 9HP; 15-JUL-92.
CC HIV: M19921; POL$NL43.
CC MEROPS: A02.001; -.
CC InterPro: IPR001969; Asprotease_AS.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR00308; Integrase_Zn.
CC InterPro: IPR009007; Pept_A_acid.
CC InterPro: IPR001995; Peptidase_A2.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR001584; Rve.
CC InterPro: IPR000477; RVse.
CC Pfam: PF00552; Integrase_1.
CC Pfam: PF02022; Integrase_Zn; 1.
CC Pfam: PF00075; rnaseH; 1.
CC Pfam: PF00665; rve; 1.
CC Pfam: PF00077; rvt; 1.
CC Pfam: PF00078; rvt; 1.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
CC PROSITE: PS50175; ASP_PROT_RETROV; 1.
CC AIDS: Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
CC CHAIN 57 155
CC ACT SITE 81 81
CC SEQUENCE 1003 AA; 113535 MW; EB9CDBA30E114A55 CRC64;

Query Match 36.8%; Score 53; DB 1; Length 1003;
Best Local Similarity 41.9%; Pred. No. 4.8;
Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITERG-----RIKKEPVHG 28
   ||:::|||||
Db 442 KALTEVPLTBEAELELAENREILKEPVHG 472

RESULT 14
```

POL_HV1MA
ID POL_HV1MA STANDARD; PRT; 1002 AA.
AC P04588; Q79582; (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 09, Last annotation update)
DT 01-NOV-1988 (Rel. 42, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL Polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (MNL isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11697;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X0415; CA28012.1; -;
CC EMBL; A07116; CA00619.1; -;
CC PDB; 1HHJ; 31-OCT-93.
CC HIV; K03456; POLSML.
CC MEROPS; A02.001; -;
DR InterPro: IPR001969; Asprotease_AS.
DR InterPro: IPR001037; Integrase_C.
DR InterPro: IPR003308; Integrase_Zn.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; naseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROT_RETROV; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferrase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 56 154 PROTEASE.
FT ACT SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113537 MW; 569A29D69AC6AC5 CRC64;
Query Match 36.1%; Score 52; DB 1; Length 1002;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 14 ITERGRITKEPVHGV 28
DB 457 LAENREITKEPVHGV 471

RESULT 15
ID POL_HV1RH STANDARD; PRT; 1002 AA.
AC P05959;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL Polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
RA Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
RA Wong-Staal F.;
RL Submitted (XX-1987) to the HIV data bank.
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA) (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17451; AAA5053.1; -;
CC HSSP; P04585; IRTN.
CC HIV; M17451; POLSRR.
CC MEROPS; A02.001; -;
DR InterPro: IPR001969; Asprotease_AS.
DR InterPro: IPR001037; Integrase_C.
DR InterPro: IPR003308; Integrase_Zn.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; naseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROT_RETROV; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferrase; RNA-directed DNA polymerase.
FT CHAIN 56 154 PROTEASE.
FT ACT SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113755 MW; 864341718E5C48C2 CRC64;
Query Match 36.1%; Score 52; DB 1; Length 1002;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 14 ITERGRITKEPVHGV 28
DB 457 LAENREITKEPVHGV 471

Tue Jul 20 07:06:41 2004

us-09-673-166b-274.rsp

Page 13

Search completed: July 20, 2004, 06:31:01
Job time : 6.5082 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 28.2295 Seconds
(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166B-274
Perfect score: 144
Sequence: 1 GRQYKANSKFIQITERGRILKEPVHGV 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL.25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_virus.*
 - 16: sp_bacteriopl.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	48.6	1310	2 Q93N27	Q93N27 clostridium
2	60	41.7	1427	15 Q9WF98	Q9WF98 human immun
3	57	39.6	1003	15 Q79792	Q79792 human immun
4	57	39.6	1003	15 P89972	P89972 human immun
5	56	38.9	322	15 Q7SKY9	Q7SKY9 human immun
6	56	38.9	323	15 Q7SK22	Q7SK22 human immun
7	56	38.9	420	15 Q8Q736	Q8Q736 human immun
8	56	38.9	420	15 Q8Q761	Q8Q761 human immun
9	56	38.9	423	15 Q8AFJ1	Q8AFJ1 human immun
10	56	38.9	524	15 Q9IDB5	Q9IDB5 human immun
11	56	38.9	560	15 Q71151	Q71151 human immun
12	56	38.9	1433	15 Q8ADE1	Q8ADE1 human immun
13	55	38.2	364	15 P90232	P90232 human immun
14	55	38.2	420	15 Q8Q784	Q8Q784 human immun
15	55	38.2	423	15 Q8JDC4	Q8JDC4 human immun
16	55	38.2	423	15 Q8AHG3	Q8AHG3 human immun

17	55	38.2	423	15 Q8AH44	Q8AH44 human immun
18	55	38.2	423	15 Q8AGZ4	Q8AGZ4 human immun
19	55	38.2	434	15 Q7ZPA1	Q7ZPA1 human immun
20	55	38.2	434	15 Q7ZP95	Q7ZP95 human immun
21	55	38.2	555	15 Q90D76	Q90D76 human immun
22	55	38.2	564	15 Q9J4V3	Q9J4V3 human immun
23	55	38.2	564	15 Q9J4V2	Q9J4V2 human immun
24	55	38.2	564	15 Q9J4V1	Q9J4V1 human immun
25	55	38.2	593	15 Q8Q876	Q8Q876 human immun
26	55	38.2	1000	15 Q8AUC5	Q8AUC5 human immun
27	55	38.2	1005	15 Q8AFB1	Q8AFB1 human immun
28	55	38.2	1429	15 Q9WF84	Q9WF84 human immun
29	54	37.5	237	15 Q9WGV3	Q9WGV3 human immun
30	54	37.5	237	15 Q9WGV7	Q9WGV7 human immun
31	54	37.5	237	15 Q9WGV5	Q9WGV5 human immun
32	54	37.5	237	15 Q9WGV4	Q9WGV4 human immun
33	54	37.5	237	15 Q9WGV6	Q9WGV6 human immun
34	54	37.5	237	15 Q9WGV3	Q9WGV3 human immun
35	54	37.5	237	15 Q9WGV4	Q9WGV4 human immun
36	54	37.5	237	15 Q9WGV2	Q9WGV2 human immun
37	54	37.5	237	15 Q9WGV5	Q9WGV5 human immun
38	54	37.5	237	15 Q9WGV6	Q9WGV6 human immun
39	54	37.5	237	15 Q9WGV3	Q9WGV3 human immun
40	54	37.5	237	15 Q9WGV4	Q9WGV4 human immun
41	54	37.5	237	15 Q9WGV6	Q9WGV6 human immun
42	54	37.5	237	15 Q9WGV3	Q9WGV3 human immun
43	54	37.5	257	15 Q9WGV4	Q9WGV4 human immun
44	54	37.5	319	15 Q7SKW0	Q7SKW0 human immun
45	54	37.5	321	15 Q9JMW5	Q9JMW5 human immun

ALIGNMENTS

RESULT 1

Q93N27 ID Q93N27 PRELIMINARY; PRT; 1310 AA.

AC Q93N27; ID 01-DEC-2001 (TREMREL. 19, Created)

DT 01-DEC-2001 (TREMREL. 19, Last sequence update)

DT 01-OCT-2003 (TREMREL. 25, Last annotation update)

DE Tetanus toxin (Fragment).

OC Clostridium tetani.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

OX NCBI_TaxId=1513;

RN [1]

RP SEQUENCE FROM N.A.

RA Shumin Z., Dianliang L.,

RT "Cloning and sequence analysis of tetanus toxin gene.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF389424; AAK72964.2; -

DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.

DR GO; GO:0008237; F:metallopeptidase activity; IEA.

DR GO; GO:0015070; F:toxin activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0009405; P:patogenesis; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR008985; ConA_like_lect_g1.

DR InterPro; IPR001064; Crystallin.

DR InterPro; IPR002160; Kunitz_legume.

DR InterPro; IPR000395; Peptidase_M27.

DR InterPro; IPR006025; Pept_M_Zn_BS.

DR Pfam; PF01742; Peptidase_M27; 1.

DR PRINTS; PR00760; BOWTOXITXIN.

DR ProDom; PD001963; Bontoxilysin; 1.

DR	PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.
FT	NON TER 1
FT	NON TER 1310
SQ	SEQUENCE 1310 AA; 150316 MW; 9EADDG314418B450 CRC64;
Query Match	48.6%; Score 70; DB 2; Length 1310;

Best Local Similarity 100.0%, Pred. No. 0.15,
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 OYKANSKFIGITE 16
| | | | | | | | | |
831 OYKANSKFIGITE 844

Db

RESULT 2
O9WF98 PRELIMINARY; PRT; 1427 AA.

AC O9WF98: 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, last sequence update)
DE 01-OCT-2003 (TEMBLrel. 25, last annotation update)
GN Gag-pol polyprotein (Gag polyprotein).
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_Taxid=11676;
RX MEDLINE=99214383; PubMed=10196340;
RA Novitsky V.A., Montano M.A., McLane M.F., Renjifo B., Vanberg F.,
Foley B.T., Ndung'u T.P., Kamran M., Makheima M.O., Marlink R.,
Essex M.;
RT "Molecular cloning and phylogenetic analysis of human immunodeficiency
virus type 1 subtype C: a set of 23 full-length clones from
Botswana";
RL J. Virol. 73:4427-4432 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C-96BWL6D14;
RA Novitsky V.A., Montano M.A., McLane M.F., Renjifo B., Vanberg F.,
Foley B.T., Ndung'u T.P., Marlink R., Essex M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
(BY SIMILARITY).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC EMBL: AF110977; AAD1753.1; --
CC PIR: S49086; S49086.
DR HSP: P03366; IHNV.
DR GO: GO:0019012; C:virion; IEA.
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004519; F:endonuclease activity; IEA.
DR GO: GO:0008907; F:integrase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0005198; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0015074; F:DNA integration; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR GO: GO:0016032; P:viral life cycle; IEA.
DR InterPro: IPR001969; Asprotease_AS.
DR InterPro: IPR000721; Gag_P24.
DR InterPro: IPR001037; Integrase_C.
DR InterPro: IPR003308; Integrase_Zn.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR000907; Pepl_A acid.
DR InterPro: IPR000071; Retrov_P17.
DR InterPro: IPR008916; Retrov_capsid_C.
DR InterPro: IPR008919; Retrov_capsid_N.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.

DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00540; Gag_P17; 1.
DR Pfam: PF00607; Gag_P24; 1.
DR Pfam: PF00552; Integrase; 1.
DR Pfam: PF02022; Integrase_Zn; 1.
DR Pfam: PF00075; rnaseh; 1.
DR Pfam: PF00655; rve; 1.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvc; 1.
DR Pfam: PF00098; zF-CCHC; 2.
DR PRINTS: PR00939; C2HCZNFINGER.
DR PRINTS: PR00234; HIVMATRIX.
DR SMART: SM00343; ZNF_C2HC; 2.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
DR PROSITE: PS0158; ZF_CCHC; 2.
DR AIDS; Aspartyl protease; Core protein; Endonuclease; Hydrolyase;
KW Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1427 AA; 160409 MW; E38F85FC6A0B369 CRC64;

Query Match 41.7%; Score 60; DB 15; Length 1427;
Best Local Similarity 73.3%; Pred. No. 5.5;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 14 ITERGRIILKEPVHGV 28
: | | | | | | | | | |
Db 882 LAENGELKEPVHGV 896

RESULT 3
O79792 PRELIMINARY; PRT; 1003 AA.

AC O79792;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, last sequence update)
DE 01-OCT-2003 (TEMBLrel. 25, last annotation update)
GN Pol protein (Fragment).
OS Gag-Pol OR Pol.
ON Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NY5;
RX MEDLINE=86259728; PubMed=3014529;
RA Willey R.L., Rutledge R.A., Dias S., Folks T., Theodore T.,
Buckler C.E., Martin M.A.;
RT "Identification of conserved and divergent domains within the envelope
gene of the acquired immunodeficiency syndrome retrovirus";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042 (1986).
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
(BY SIMILARITY).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC EMBL: M38431; AAB04037.1; --
CC PIR: A47330; A47330.
DR PIR: B47330; B47330.
DR PIR: C47330; C47330.
DR PIR: D47330; D47330.
DR PIR: E47330; E47330.
DR PIR: F47330; F47330.
DR PIR: S32058; S32058.
DR PIR: S32132; S32132.
DR PIR: S32140; S32140.
DR HSP: P12497; SHVP.
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004519; F:endonuclease activity; IEA.
DR GO: GO:0008907; F:integrase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004523; F:ribonuclease H activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.

```
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00075; rnaesh; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Endonuclease; Hydrolyase; Nucleotidyltransferase;
KW Polyprotein; Protease; RNA-directed DNA polymerase; Transferase.
FT NON_TER
SQ SEQUENCE 1003 AA; 113729 MW; E24A73675C1651A1 CRC64;

Query Match 39.6%; Score 57; DB 15; Length 1003;
Best Local Similarity 43.8%; Pred. No. 11;
Matches 14; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

QY 5 KANSKFIGITRG-----RIIKEPVHG 28
DB 441 KALTEVPLTEAELELAENREILKEPVHG 472

RESULT 4
P89972 PRELIMINARY; PRT; 1003 AA.
AC P89972;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol polyprotein (Fragment).
GN GAG-POL OR POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Ghosh S., Ghosh E., Shaw G.M., Hahn B.;
RT "Human Immunodeficiency Virus type 1 complete genomic sequence.";
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MacInnes K.A.;
RT Submitted (FEB-1995) to the EMBL/Genbank/DBJ databases.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL; U21135; AAC32294.1; -.
DR PIR; A47330; A47330.
DR PIR; B47330; B47330.
DR PIR; C47330; C47330.
DR PIR; D47330; D47330.
DR PIR; E47330; E47330.
DR PIR; F47330; F47330.
DR PIR; S32132; S32132.
DR HSSP; P04585; 1RTH.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
```

```
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA recombination; IEA.
DR GO; GO:0006310; P:DNA integration; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00075; rnaesh; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Endonuclease; Hydrolyase; Nucleotidyltransferase;
KW Polyprotein; Protease; RNA-directed DNA polymerase; Transferase.
FT NON_TER
SQ SEQUENCE 1003 AA; 113833 MW; 29A57DAF295477A9 CRC64;

Query Match 39.6%; Score 57; DB 15; Length 1003;
Best Local Similarity 48.4%; Pred. No. 11;
Matches 15; Conservative 2; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIIKEPVHG 28
DB 442 KALTEIIPITBEAELELAENREILKEPVHG 472

RESULT 5
Q7SKY9 PRELIMINARY; PRT; 322 AA.
AC Q7SKY9;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3999-7303.4;
RA Gonzales M.J., Johnston E., Dupnik K.M., Imamichi T., Shafer R.W.;
RT "Coincidence of Reverse Transcriptase Inhibitor Resistance Mutations
RT Detected by Population-based Sequencing.";
RL EMBL; AY351751; AAQ18859.1; -.
DR EMBL; AY351751; AAQ18859.1; -.
DR RNA-directed DNA polymerase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 322 AA; 37354 MW; 095623DBAC3D0D5 CRC64;

Query Match 38.9%; Score 56; DB 15; Length 322;
Best Local Similarity 45.2%; Pred. No. 4.4;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIIKEPVHG 28
||:|:|:|
```



```

Q8AFL1
ID Q8AFL1 PRELIMINARY; PRT; 423 AA.
AC Q8AFL1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=lv904;
RA Brindeiro P.A., Brindeiro R.M., Mortensen C., Hertogs K., de Vroey V.,
RA Tannu N.P.M., Sion F.S., de Sa C.A.M., Machado D.M., Succi R.C.M.,
RA Tannu A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY145784; AA012458.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR009007; Peptidase A2.
DR InterPro; IPR000477; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 423 AA; 48130 MW; 32F367BF19B068F1 CRC64;

Query Match
Best Local Similarity 45.2%; Score 56; DB 15; Length 423;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIIKEPVHGV 28
Db 386 KALTEVISITBEAELELAENREILKEPVHGV 416

RESULT 10
Q9IDB5 PRELIMINARY; PRT; 524 AA.
AC Q9IDB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Pol (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=99FR-WP1321;
RX MEDLINE=20514217; PubMed=11060045;
RX Verge L., Peters M., Mpoudi-Ekoule E., Bourgeois A., Liegeois F.,
RX Toure-Kane C., Moup S., Mulanga-Kabeya C., Saman E., Jourdan J.,
RX Reyes J., Delaporte E.;
RT "Genetic Diversity of Protease and Reverse Transcriptase Sequences in
RT non-subtype-b human immunodeficiency virus type 1 strains: evidence of
RT many minor drug resistance mutations in treatment-naïve patients.";
RL J. Clin. Microbiol. 38:3919-3925(2000).
CC -1- STIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPTIN FAMILY.
DR EMBL; AJ287045; CAC03087.1; -.
DR PIR; A47330; A47330.
DR PIR; C47330; C47330.

```

```

DR PIR; D47330; D47330.
DR PIR; F47330; F47330.
DR HSSP; P04585; 1RTH.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:peptidase activity; IEA.
DR GO; GO:0003964; F:RNA binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 524 AA; 60460 MW; 8780BEA11696EB7C CRC64;

Query Match
Best Local Similarity 45.2%; Score 56; DB 15; Length 524;
Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 6 KANSKFIGITRG-----RIIKEPVHGV 28
Db 389 KALTEVISITBEAELELAENREILKEPVHGV 419

RESULT 11
Q71151 PRELIMINARY; PRT; 560 AA.
AC Q71151;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE Pol (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VE4;
RX MEDLINE=96093896; PubMed=7576917;
RX Quinones-Mateu M.E., Holguin A., Dopazo J., Rota T.R., Domingo E.;
RX "Molecular characterization of human immunodeficiency virus type 1
RX isolates from Venezuela.";
RL AIDS Res. Hum. Retroviruses 11:605-616(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VE4;
RX MEDLINE=97000986; PubMed=8844016;
RX Quinones-Mateu M.E., Holguin A., Dopazo J., Najera I., Domingo E.;
RT "Point mutant frequencies in the pol gene of human immunodeficiency
RT virus type 1 are two- to threefold lower than those of env.";
RL AIDS Res. Hum. Retroviruses 12:1117-1128(1996).
DR EMBL; U16771; AAC55688.1; -.
DR PIR; A47330; A47330.
DR PIR; B47330; B47330.
DR PIR; C47330; C47330.
DR PIR; D47330; D47330.
DR PIR; F47330; F47330.
DR PIR; S32132; S32132.
DR PIR; S32140; S32140.
DR HSSP; P03366; 1HRH.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.

```

DR GO:0016740; F:transferase activity; IEA.
 DR GO:0006278; P:RNA dependent DNA replication; IEA.
 DR InterPro: IPR002156; RNaseH.
 DR InterPro: IPR000477; RVase.
 DR Pfam: PF00075; rnasel; 1.
 DR Pfam: PF00078; rvt; 1.
 DR RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 560
 SQ SEQUENCE 560 AA; 64354 MW; 8C29E9B280090C8F CRC64;
 Query Match 38.9%; Score 56; DB 15; Length 560;
 Best Local Similarity 45.2%; Pred. No. 8.1;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 6 KANSKPIGITERG-----RIKKEPVHGV 28
 Db 287 KALTEVISTTEAELELANRELKEPVHGV 317
 RESULT 12
 Q8ADE1 PRELIMINARY; PRT; 1433 AA.
 AC Q8ADE1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Gag-pol fusion polypeptide (Fragment).
 GN GAG-POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=98UC57143;
 RX MEDLINE=20564795; PubMed=11112486;
 RA Salminen M.O., Ehrenberg P.K., Mascola J.R., Dayhoff D.E., Merling R., Blake B., Linder M., Hegerich S., Polonis V.R., Birx D.L., Robb M.L., McCutchan F.E., Michael N.L.;
 RA "Construction and biological characterization of infectious molecular clones of HIV-1 subtypes B and E (CRF01_AE) generated by the polymerase chain reaction";
 RT Virology 278:103-110(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=98UC57143;
 RX MEDLINE=22375625; PubMed=12487816;
 RA Harris M.E., Serwadda D., Sewankambo N., Wabwire F., Kim B., Kigozi G., Kiwanuka N., Phillips J.B., Meehan M., Lutalo T., Lane J.R., Merling R., Gray R., Wawer M., Birx D.L., Robb M.L., McCutchan F.E.;
 RA "Among 46 Near Full Length HIV Type 1 Genome Sequences from Rakai District, Uganda, Subtype D and AD1 Recombinants Predominate";
 RT AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
 RL EMBL: AF484514; AAN73763.1;
 DR GO:00004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO:0003677; F:DNA binding; IEA.
 DR GO:0008907; F:integrase activity; IEA.
 DR GO:0004523; F:ribonuclease H activity; IEA.
 DR GO:0003723; F:RNA binding; IEA.
 DR GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO:0005198; F:structural molecule activity; IEA.
 DR GO:0008270; F:zinc ion binding; IEA.
 DR GO:0015074; P:DNA integration; IEA.
 DR GO:0006310; P:DNA recombination; IEA.
 DR GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO:0006278; P:RNA dependent DNA replication; IEA.
 DR GO:0016032; P:virial life cycle; IEA.
 DR InterPro: IPR001969; Aspartase_AS.
 DR InterPro: IPR000721; Gag_p24.
 DR InterPro: IPR001037; Integrase_C.
 DR InterPro: IPR003308; Integrase_Zn.
 DR InterPro: IPR001995; Peptidase_A2.

DR InterPro: IPR009007; Pept_A acid.
 DR InterPro: IPR000071; Retrovir_p17.
 DR InterPro: IPR008916; Retrov_capsid_C.
 DR InterPro: IPR008919; Retrov_capsid_N.
 DR InterPro: IPR002156; RNaseH.
 DR InterPro: IPR001584; Rve.
 DR InterPro: IPR000477; RVase.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF00540; Gag_p17; 1.
 DR Pfam: PF00607; Gag_p24; 1.
 DR Pfam: PF00552; Integrase; 1.
 DR Pfam: PF02022; Integrase_Zn; 1.
 DR Pfam: PF00075; rnasel; 1.
 DR Pfam: PF00655; rve; 1.
 DR Pfam: PF00077; rvt; 1.
 DR Pfam: PF00078; rvt; 1.
 DR Pfam: PF00098; zf_CCHC; 2.
 DR PRINTS: PR00939; C2HCZNFINGER.
 DR PRINTS: PR00234; HIVMATRIX.
 DR SMART: SM00343; ZNF_C2HC; 2.
 DR PROSITE: PS00141; ASP_PROTASE; 1.
 DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
 DR PROSITE: PS50158; ZF_CCHC; 2.
 KW Polypeptide.
 FT NON_TER 1
 FT NON_TER 161645
 SQ SEQUENCE 1433 AA; 161645 MW; EB213A9BCB7D583 CRC64;
 Query Match 38.9%; Score 56; DB 15; Length 1433;
 Best Local Similarity 45.2%; Pred. No. 22;
 Matches 14; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 6 KANSKPIGITERG-----RIKKEPVHGV 28
 Db 872 KALTEVISTTEAELELANRELKEPVHGV 902
 RESULT 13
 P90232 PRELIMINARY; PRT; 364 AA.
 AC P90232;
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE POL polypeptide (Fragment).
 GN POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IND 4;
 RX MEDLINE=97104208; PubMed=8948366;
 RA Soto-Ramirez L.B., Tripathy S., Remjifo B., Essex M.;
 RA "HIV-1 pol sequences from India fit distinct subtype pattern";
 RT J. Acquir. Immune Defic. Syndr. Hum. Retrovir. 13:293-307(1996).
 RL EMBL: U31242; AAB40987.1;
 DR HSSP: P03366; IDLO.
 DR GO:0003723; F:RNA binding; IEA.
 DR GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO:0016740; F:transferase activity; IEA.
 DR GO:0006278; P:RNA dependent DNA replication; IEA.
 DR InterPro: IPR000477; RVase.
 DR Pfam: PF00078; rvt; 1.
 KW Polypeptide; RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 364
 FT CHAIN 8
 FT CHAIN 364
 FT NON_TER 42031
 SQ SEQUENCE 364 AA; 42031 MW; F4F0479B008590AF CRC64;
 Query Match 38.2%; Score 55; DB 15; Length 364;
 Best Local Similarity 40.6%; Pred. No. 7.2;
 Matches 13; Conservative 4; Mismatches 7; Indels 8; Gaps 1;

QY 5 IKANSKFIGITRG-----RIKEPVHGV 28
 ID : : : : :
 DB 293 KALTEVVSILTEBALELAENREILKEPVHGV 324

RESULT 14
 O80784 PRELIMINARY; PRT; 420 AA.
 ID O80784;
 AC O80784;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Pol protein (Fragment).
 GN POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RP [1]
 RC SEQUENCE FROM N.A.
 RA Eshleman S.H., Becker-Pergola G., Deseve M., Macna M., Guay L.A.,
 RA Cunningham S., Musoke P., Mmlro F., Jackson J.B.;
 RT "Impact of HIV-1 subtype on women receiving single dose NVP
 RT prophylaxis to prevent HIV-1 vertical transmission (HIVNET012).";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 DR EMBL; AF388099; AAL84041.1; -.
 DR HSSP; P03366; 2HM1.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
 DR InterPro: IPR001969; Asprotease AS.
 DR InterPro: IPR009007; Peptidase A2.
 DR InterPro: IPR00477; RVTse.
 DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvt; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
 KW Transierase.
 KM NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 420 AA; 48035 MW; E871192B0C277E80 CRC64;

Query Match 38.2%; Score 55; DB 15; Length 420;
 Best Local Similarity 41.9%; Pred. No. 8.4;
 Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGITRG-----RIKEPVHGV 28
 ID : : : : :
 DB 386 KALTEVVSILTEBALELAENREILKEPVHGV 416

RESULT 15
 O8JDC4 PRELIMINARY; PRT; 423 AA.
 ID O8JDC4;
 AC O8JDC4;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Pol protein (Fragment).
 GN POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=148154F;
 RA Macna M.R., Eshleman S.H.;
 RT "Impact of maternal HIV-1 subtype on Ugandan infants (HIVNET 012).";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF410250; AAM02560.1; -.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
 DR InterPro: IPR001969; Asprotease AS.
 DR InterPro: IPR009007; Peptidase A2.
 DR InterPro: IPR00477; RVTse.
 DR Pfam; PF00077; rvp; 1.
 DR Pfam; PF00078; rvt; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 KW RNA-directed DNA polymerase; Transferase.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 423 AA; 48265 MW; EEF645A37D43C405 CRC64;

Query Match 38.2%; Score 55; DB 15; Length 423;
 Best Local Similarity 41.9%; Pred. No. 8.4;
 Matches 13; Conservative 4; Mismatches 6; Indels 8; Gaps 1;
 QY 6 KANSKFIGITRG-----RIKEPVHGV 28
 ID : : : : :
 DB 386 KALTEVVSILTEBALELAENREILKEPVHGV 416

Search completed: July 20, 2004, 06:33:16
 Job time : 29.2295 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 41.7705 Seconds

(without alignments)
189,400 Million cell updates/sec

Title: US-09-673-166B-274

Perfect score: 144
Sequence: 1 GRQYTKANSKFTGTERGRILKEPVHGV 28Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111.5	77.4	25	4	AA884738 Aab84738 Amino aci
2	97	67.4	28	2	AA53552 Aay53552 Lipopepti
3	86	59.7	29	2	AA53551 Aay53551 Lipopepti
4	81	56.2	29	5	ABB79182 Human cyt
5	79	54.9	216	3	AA92665 MUC-1 ana
6	78.5	54.5	31	3	AA82632 Tetanus t
7	78	54.2	750	3	AA92639 Mutant hu
8	77	53.5	37	2	AA83561 IGE CH4 r
9	76	52.8	29	2	AA83561 IGE CH4 r
10	76	52.8	37	2	AA83561 IGE CH4 r
11	76	52.8	47	2	AA83561 IGE CH4 r
12	76	52.8	47	2	AA83561 IGE CH4 r
13	75	52.1	25	3	AA92652 Aay92652 PSMPep009
14	75	52.1	25	3	AA92652 Aay92652 PSMPep009
15	75	52.1	31	5	ABB79179 Human cyt
16	75	52.1	50	2	AAW06132 Anti-choi
17	75	52.1	50	2	AAW06132 Anti-choi
18	75	52.1	109	4	AAH20145 Growth di
19	75	52.1	158	2	AAH81329 TNF2-4, a
20	75	52.1	158	2	AAH81330 TNF2-5, a
21	75	52.1	158	5	ABB07281 Human TNF
22	75	52.1	158	5	ABB07281 Human TNF
23	75	52.1	693	3	AA92649 Mutant hu
24	75	52.1	750	3	AA92645 Mutant hu
25	75	52.1	750	3	AA92627 Mutant hu

26	75	52.1	750	3	AA92640 Mutant hu
27	75	52.1	750	3	AA92630 Mutant hu
28	75	52.1	750	3	AA92646 Mutant hu
29	75	52.1	750	3	AA92641 Mutant hu
30	75	52.1	28	5	AAU11422 Synthetic
31	74	51.4	31	5	ABB79180 Human cyt
32	74	51.4	32	3	AA82636 Tetanus t
33	74	51.4	126	3	AA82636 Tetanus t
34	74	51.4	137	3	AA82634 Tetanus t
35	74	51.4	514	6	AA030491 Human TNF
36	74	51.4	514	6	AA030490 Human TNF
37	74	51.4	514	6	AA030490 Human TNF
38	74	51.4	517	6	AA030492 Human TNF
39	73	50.7	25	3	AA92650 PSMPep007
40	73	50.7	31	2	AAW06129 Anti-choi
41	73	50.7	31	2	AAW06129 Anti-choi
42	73	50.7	693	3	AA92647 Mutant hu
43	73	50.7	750	3	AA92628 Mutant hu
44	73	50.7	750	3	AA92644 Mutant hu
45	72	50.0	17	2	AA62692 Helper T

ALIGNMENTS

RESULT 1	AA884738	standard; peptide; 25 AA.
ID	AA884738	
AC	AA884738;	
DT	17-SEP-2001	(first entry)
DE	Amino acid sequence of lipopeptide M2-K.	
XX	Lipid-tailed protein; mucosal membrane; immune system; lipoprotein;	
KW	B cell response; T cell response; intranasal immunisation;	
KW	sublingual immunisation.	
XX	Synthetic.	
OS		
FH	Key	Location/Qualifiers
FT	Modified-site	1
FT		/note= "amidated residue with palmitoyl lipid residue with 16 carbon atoms attached via an epsilon bond"
FT		
PN	MO200141797-AZ.	
XX		
PD	14-JUN-2001.	
XX		
PF	08-DEC-2000; 2000WO-EP012794.	
XX		
PR	09-DEC-1999; 99US-0169952P.	
XX		
PA	(INSP) INST PASTEUR.	
PA	(UYLI-) UNIV LILLE 2.	
XX		
PI	Druihe P, Gras-Masse H, Benmohamed L;	
XX		
DR	WPI; 2001-451597/48.	
XX		
PT	Inducing systemic immune response comprises mucosal administration of lipid-tailed proteins or lipopeptide in absence of adjuvants.	
XX		
PS	Claim 22; Page 26; 34pp; English.	
XX		
CC	The present sequence represents a lipid-tailed protein. The lipid moiety is able to cross the mucosal membrane, and deliver an antigen to the immune system. The lipoprotein is administered to a mucosal membrane to induce an immune response. The lipoprotein induces a B cell and/or a T cell response. The lipoprotein is used to provide intranasal or sublingual immunization	
CC		
CC		
XX		

SQ Sequence 25 AA; 77.4%; Score 111.5; DB 4; Length 25;
 Query Match Best Local Similarity 96.0%; Pred. No. 5.5e-11;
 Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GROYT-KANSKFIGITERGRILKEP 24
 Db 1 GROYIKRANSKFIGITERGRILKEP 25

RESULT 2
 AAY53552
 ID AAY53552 standard; protein; 28 AA.
 AC AAY53552;
 XX
 XX 18-JAN-2000 (first entry)
 DE Lipopeptide #3.
 XX
 XX Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
 KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
 KM human immunodeficiency virus; hepatitis B virus; papilloma virus;
 KM melanoma; malaria; parasite.
 OS Synthetic.
 OS Homo sapiens.
 FH Key
 FT Modified-site 1 Location/Qualifiers
 FT 1 /note="contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"
 PN FR2776926-A1.
 PD 08-OCT-1999.
 PF 07-APR-1998; 98FR-00004323.
 XX
 XX 07-APR-1998; 98FR-00004323.
 PR (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA (INSP) INST PASTEUR LILLE.
 PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
 DR WPI; 1999-583113/50.
 DX
 PT New lipopeptide containing lipid regions and two epitopes, all separated by peptide spacers that impart hydrophilicity, useful in vaccines.
 PS Example 1; Page 9; 35pp; French.
 XX
 XX The invention relates to the generation of a lipopeptide comprising at least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL) epitope and at least one lipid residue with (i) the epitopes and lipid portion and (ii) the epitopes, being separated independently by peptide spacers. These spacers comprise sequences of amino acids which carry an overall electrical charge in neutral media to ensure that the lipopeptide is hydrophilic. This peptide represents an example of them lipopeptide of the invention. It contains 2 lipid residues attached at the N-terminal end of the molecule and 2 epitopes separated by the spacer residues Ser-Ser and Ala-Ala-Ala. Peptides AAY53301-Y53549 represent peptide epitopes used in the generation of the lipopeptides. These are used in therapeutic or prophylactic compositions and vaccines to induce specific immune responses against human immunodeficiency, hepatitis B or papilloma viruses; p53 of melanoma or the malaria parasite

SQ Sequence 28 AA; 67.4%; Score 97; DB 2; Length 28;

Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GROYIKANSKFIGITERGR 19
 Db 1 GROYIKANSKFIGITERGR 19

RESULT 3
 AAY53551
 ID AAY53551 standard; protein; 29 AA.
 AC AAY53551;
 XX
 XX 18-JAN-2000 (first entry)
 DE Lipopeptide #2.
 XX
 XX Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
 KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
 KM human immunodeficiency virus; hepatitis B virus; papilloma virus;
 KM melanoma; malaria; parasite.
 OS Synthetic.
 OS Homo sapiens.
 FH Key
 FT Modified-site 1 Location/Qualifiers
 FT 1 /note="contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"
 PN FR2776926-A1.
 PD 08-OCT-1999.
 PF 07-APR-1998; 98FR-00004323.
 XX
 XX 07-APR-1998; 98FR-00004323.
 PR (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA (INSP) INST PASTEUR LILLE.
 PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
 DR WPI; 1999-583113/50.
 DX
 PT New lipopeptide containing lipid regions and two epitopes, all separated by peptide spacers that impart hydrophilicity, useful in vaccines.
 PS Example 1; Page 9; 35pp; French.
 XX
 XX The invention relates to the generation of a lipopeptide comprising at least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL) epitope and at least one lipid residue with (i) the epitopes and lipid portion and (ii) the epitopes, being separated independently by peptide spacers. These spacers comprise sequences of amino acids which carry an overall electrical charge in neutral media to ensure that the lipopeptide is hydrophilic. This peptide represents an example of them lipopeptide of the invention. It contains 2 lipid residues attached at the N-terminal end of the molecule and 2 epitopes separated by the spacer residues Ser-Ser and Ala-Ala-Ala. Peptides AAY53301-Y53549 represent peptide epitopes used in the generation of the lipopeptides. These are used in therapeutic or prophylactic compositions and vaccines to induce specific immune responses against human immunodeficiency, hepatitis B or papilloma viruses; p53 of melanoma or the malaria parasite

SQ Sequence 29 AA; 59.7%; Score 86; DB 2; Length 29;

Query Match Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 OYIKANSKFIGITERGR 19
DB 4 OYIKANSKFIGITERGR 20

RESULT 4
ID ABB79182
AC ABB79182; standard; peptide; 29 AA.
XX ABB79182;
XX 07-AUG-2002 (first entry)
XX
DE Human cytomegalovirus PPI50 related vaccine peptide SEQ ID NO:8.
XX
XX Human cytomegalovirus PPI50; HCMV; human leukocyte antigen; HLA; CTL;
KW cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
KW immunogenic; virucide; immune response.
XX
OS Human cytomegalovirus.
XX
PN WO200234769-A2.
XX
XX 02-MAY-2002.
XX
XX 22-OCT-2001; 2001MO-US032589.
XX
XX 20-OCT-2000; 2000US-0241944P.
XX
XX (CITY) CITY OF HOPE.
XX
XX Diamond DJ;
XX
DR WPI; 2002-471432/50.
XX
PT New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
PT cytomegalovirus PPI50, useful for creating vaccines against
XX cytomegalovirus infection.
XX
PS Disclosure; Page 14; 28pp; English.
XX
XX The present invention describes a peptide (I) which is an immunogenic
CC epitope recognised by CD8+ class I major histocompatibility complex (MHC)
CC restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent
CC cytomegalovirus (HCMV) infection. Also described are: (1) a vaccine (II)
CC against human cytomegalovirus (CMV), which comprises (I); (2) a cellular
CC vaccine (III) against human CMV comprises antigen presenting cells which
CC present (I); (3) a recombinant viral vector (IV) which contains a gene
CC encoding (I); and (4) an immunological reagent (V) which comprises (I).
CC (I) has virucide activity. (II) or (III) are useful for modulating the
CC immune response to HCMV infection, and for vaccinating a mammal in need
CC of it against HCMV. (I) is useful for eliciting a cellular immune against
CC HCMV by normal and immunodeficient subjects, for creating efficient
CC vaccines against HCMV, and in immunological methods to detect PPI50-
CC reactive CTL in a patient or a sample from the patient. The present
CC sequence represents a HCMV PPI50 related vaccine peptide, which is given
CC in the exemplification of the present invention
XX
XX
SQ Sequence 29 AA;
QY
Query Match 56.2%; Score 81; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 GROVIKANSKFIGITE 16
14 GROVIKANSKFIGITE 29

RESULT 5
AAV92665
ID AAV92665 standard; peptide; 216 AA.
XX

AC AAV92665;
XX
XX 10-AUG-2000 (first entry)
XX
DE MUC-1 analogue containing foreign epitopes.
XX
XX Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
KW cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
KW cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Peptide 61..75
FT /label= P2
FT 136..156
FT Peptide /label= P30
FT /note= "q"
XX
XX
PN WO20020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK000525.
XX
XX 05-OCT-1998; 98DK-00001261.
XX
XX 20-OCT-1998; 98US-0105011P.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
XX
PS Example 4; Page; 220pp; English.
XX
XX This is an immunogenized MUC-1 analogue containing foreign epitopes P2
CC and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate specific
CC membrane antigen (hPSM) can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms (see features table). 10 regions
CC suitable for the insertion of foreign T helper epitopes were identified.
CC The method is used for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
CC and/or fibroblast growth factor 8b (FGF8b). The method comprises
CC effecting simultaneous presentation by antigen producing cells (APCs) of
CC the animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively. Note: This sequence does not appear in
CC the specification. It was made using the mucin repeat sequence
CC (AAV92664), P2 and P30 (AAV92625-26), which appear on pages 220, 213 and
CC 214 respectively, of the specification
XX
XX
SQ Sequence 216 AA;
QY
Query Match 54.9%; Score 79; DB 3; Length 216;
Best Local Similarity 65.5%; Pred. No. 0.00016;
Matches 19; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
DB 2 ROYIKANSKFIGITE--RGRIKPEVHGV 28
ROYIKANSKFIGITE--RGRIKPEVHGV 28

Db 60 ROYIKANSKFIGITELPAPGSTAPPAGHV 89

RESULT 6
AAV82632
ID AAV82632 standard; peptide; 31 AA.
XX AAV82632;
XX
XX 07-AUG-2000 (first entry)
XX
XX Tetanus toxoid T cell epitope and Der pii B cell epitope peptide.
XX
XX T cell epitope; B cell epitope; allergy; allergen; antigenic;
XX
XX antiallergic; antihastmatic; antiinflammatory; dermatological;
XX
XX immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
XX
XX atopic dermatitis; acute urticaria; chronic urticaria;
XX
XX gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
XX
XX anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
XX Dermatophagoides pteronyssinus.
XX
XX Clostridium tetani.
XX
XX Synthetic.
XX
XX MO200006694-A2.
XX
XX 10-FEB-2000.
XX
XX 20-JUL-1999; 99WO-BE000092.
XX
XX 30-JUL-1998; 98EP-00870167.
XX
XX (UNIO) UCB SA.
XX
XX Saint-Remy J. Jacquemin M;
XX
XX WPI; 2000-422470/36.
XX
XX New compound for prevention and treatment of allergies comprises at least
XX
XX one allergen antigenic determinant recognized by a B cell and at least
XX
XX one antigenic determinant which does not trigger T cell activation.
XX
XX
XX Claim 8; Page 35; 50pp; English.

The present invention describes a compound (I) for the prevention and/or treatment of allergy. The compound comprises at least one allergen antigenic determinant (i) recognised by a B cell or an antibody secreted by a B cell of a non-atopic individual and at least one antigenic determinant (ii) different from the allergen that triggers T cell activation. (I) has antiallergic, antihastmatic, antiinflammatory, dermatological and immunosuppressive activities, and can be used in a vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to treat and/or prevent allergies or a disease of allergic origin, especially hypersensitivities. These include rhinitis, sinusitis, bronchial asthma, atopic dermatitis, some forms of acute and chronic urticaria, gastro-intestinal syndromes associated with the ingestion of food allergens, oro-pharyngeal syndrome, anaphylactic reactions associated with drug hypersensitivities and/or a mixture of these. The use of (I) in the treatment of allergic conditions avoids the need for drug treatment, which often causes undesirable side-effects. Also, prior art drug therapies alleviate symptoms, but do not influence their causes, however (I) actually combats the cause of an allergic reaction. The present sequence represents a specifically claimed compound peptide sequence from the present invention

Sequence 31 AA;
54.5%; Score 78.5; DB 3; Length 31;
Best Local Similarity 60.0%; Pred. No.2e-05;
Matches 18; Conservative 2; Mismatches 5; Indels 5; Gaps 1

3 QYIKANSKFIGITTEG-----RIIKPEVHG 27
|||||
:|:|

Db 1 QYIKANSKFIGITELGHEIKKVLVPGCHG 30

RESULT 7
AA92639
ID AA92639 standard; protein; 750 AA.
XX
XX AA92639;
AC
AD 10-AUG-2000 (first entry)
DE Mutant human prostate specific membrane antigen construct, hPSM5.1.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KM prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 21..41
FT /label= P30
FT /note= "foreign epitope"
FT 305..319
FT Peptide /label= P2
FT /note= "foreign epitope"
FT
FN WO200020027-A2.
PN
PD 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK000525.
PP
XX 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
XX
XX (MEBT-) M & E BIOTECH AS.
PA
PI Steinaa I., Mouritsen S., Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI, 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
XX Example 1; Page; 220pp; English.
XX
XX AA92639-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX The immunogenic analogues of PSM can be used in the claimed method as an
XX autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX binding regions and cysteine residues involved in disulfide bonds are
XX preserved in the immunogenized forms. The method is used for inducing
XX immune responses against weakly immunogenic cell-associated peptide
XX antigens (PA) such as those associated with cancers (self-proteins), e.g. or
XX human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX simultaneous presentation by antigen producing cells (APCs) of the
XX animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX cell derived from the PA and/or at least 1 B-cell group derived from the
XX cell-associated PA; and (2) at least 1 first T helper cell group which is
XX foreign to the animal. Analogues of human PSM, human Her2 and
XX human/murine FGF8b comprising a substantial part of all known and
XX predicted CTL and B-cell epitopes of the respective PA and including at
XX least one foreign T helper epitope are also claimed. The method is used
XX to treat prostate, prostate/breast or breast cancer when the PA is human
XX PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX from the wild type human PSM (AA92639), which appears on pages 184-187
XX of the specification
XX
XX Sequence 750 AA;
SQ

Query Match 54.2%; Score 78; DB 3; Length 750;
 Best Local Similarity 73.9%; Pred. No. 0.001;
 Matches 17; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 2 ROYIKANSKFIGITERGRILKEP 24
 :|||||
 Db 304 KOYIKANSKFIGITELRGSLKVP 326

RESULT 8
 AAR65389
 ID AAR65389 standard; peptide; 37 AA.
 XX
 AC AAR65389;
 DT 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)
 XX
 DE Universal immunostimulator having GG spacers.
 XX
 KM Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW tetanus toxin.
 XX
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Domain 3..19
 FT /note= "tetanus toxin helper T cell epitope"
 FT Domain 22..37
 FT /note= "invasin domain"
 XX
 PN WO9425060-A1.
 XX
 PD 10-NOV-1994.
 PD
 PF 28-APR-1994; 94WO-US004832.
 XX
 PR 27-APR-1993; 93US-00057166.
 PR 14-APR-1994; 94US-00229275.
 XX
 PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 PI Ladd AE, Wang CY, Zamb T;
 DR WPI; 1994-357910/44.
 XX
 PT Immunogenic luteinizing hormone releasing hormone peptide(s) - that
 PT suppresses LHRH activity in males and females.
 XX
 PS Disclosure; Page 95; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasin protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 CC domains and between the immune stimulator and hapten components. When the
 CC hapten is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence is an example of a
 CC -GG-Th-GG-invasin immune stimulator to which a hapten can be bonded.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX
 SQ Sequence 37 AA;

Query Match 53.5%; Score 77; DB 2; Length 37;
 Best Local Similarity 76.2%; Pred. No. 4.3e-05;

Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ROYIKANSKFIGITERGRILK 22
 :|||||
 Db 4 KOYIKANSKFIGITELGTRAK 24

RESULT 9
 AAR83561
 ID AAR83561 standard; peptide; 29 AA.
 XX
 AC AAR83561;
 DT 13-JUN-1996 (first entry)
 DT
 DE IGE CH4 region contg. peptide immunogen for treating allergies.
 XX
 KM IGE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 KW vaccine; allergy; antibody; constant heavy chain.
 XX
 OS Synthetic.
 FH WO9526365-A1.
 PN
 PD 05-OCT-1995.
 PD
 PF 24-MAR-1995; 95WO-US003741.
 XX
 PR 28-MAR-1994; 94US-00218461.
 PR 25-OCT-1994; 94US-00328912.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 DR WPI; 1995-351297/45.
 XX
 PT Synthetic peptide-based immunogen contg. IGE CH4 peptide and helper T
 PT cell epitope - useful for eliciting antibody produ. for allergy
 PT treatment.
 XX
 PS Claim 5; Page 68-69; 87pp; English.
 XX
 CC AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are
 CC useful in vaccines for treating allergic reactions. In the immunogens, an
 CC IGE CH4 peptide is attached C-terminally to a series of amino acids
 CC including a helper T cell epitope. The immunogen may also opt. contain a
 CC fatty acid or fatty acid derivative, an invasin domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site in
 CC human IGE heavy chain (the CH4 domain peptide) which inhibit mast cell
 CC activation and reduce allergen-induced IGE prodn. The immunogens may be
 CC used in either a radially branching multimeric form or a linearly
 CC arranged monomeric form
 XX
 SQ Sequence 29 AA;

Query Match 52.8%; Score 76; DB 2; Length 29;
 Best Local Similarity 88.2%; Pred. No. 4.7e-05;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ROYIKANSKFIGITERG 18
 :|||||
 Db 2 KOYIKANSKFIGITELG 18

RESULT 10
 AAR65383
 ID AAR65383 standard; peptide; 37 AA.
 XX
 AC AAR65383;
 DT 25-MAR-2003 (revised)
 DT 21-SEP-1995 (first entry)

```

XX Universal immunostimulator having GG spacers.
XX Helper T cell epitope; universal immune stimulator; invasin; hapten;
XX tetanus toxin.
XX Synthetic.
XX Key Location/Qualifiers
XX FT 1..16
XX FT Domain /note= "invasin domain"
XX FT 19..35
XX FT Domain /note= "tetanus toxin helper T cell epitope"
XX WO9425060-A1.
XX PD 10-NOV-1994.
XX PF 28-APR-1994; 94WO-US004832.
XX PR 27-APR-1993; 93US-00057166.
XX PR 14-APR-1994; 94US-00229275.
XX PA (LADD/) LADD A E.
XX PA (WANG/) WANG C Y.
XX PA (ZAMB/) ZAMB T.
XX PI Ladd AE, Wang CY, Zamb T;
XX PI MPI, 1994-357910/44.
XX PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
XX PT suppress LHRH activity in males and females.
XX PS Disclosure; Page 95; 213pp; English.
XX CC Synthetic immunogenic peptides are provided in which a universal immune
XX CC stimulator is linked to a peptide or protein hapten containing B cell
XX CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX CC potent immune responses to the coupled peptide or protein. The stimulator
XX CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
XX CC immune response to the coupled peptide in members of a heterogeneous
XX CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
XX CC sequence from the invasin protein of Yersinia. Spacer amino acid
XX CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
XX CC domains and between the immune stimulator and hapten components. When the
XX CC hapten is LHRH, then optionally the invasin domain can be omitted from
XX CC the immune stimulator component. The present sequence is an example of an
XX CC invasin-GG-Th-GG- immune stimulator to which a hapten can be bonded.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 37 AA;
XX Query Match 52.8%; Score 76; DB 2; Length 37;
XX Best Local Similarity 88.2%; Pred. No. 6.2e-05;
XX Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RQYIKANSKFTIGTERG 18
Db :|||||
20 KQYIKANSKFTIGTELG 36

```

```

RESULT 11
AAR62723
ID AAR62723 standard; peptide; 47 AA.
XX
XX AAR62723;
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 17-SEP-1995 (first entry)
XX
XX LHRH-containing immunogenic peptide.
XX
XX

```

```

XX Helper T cell epitope; universal immune stimulator; invasin; hapten;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
XX Synthetic.
XX Key Location/Qualifiers
XX FT 1..16
XX FT Domain /note= "invasin domain"
XX FT 19..35
XX FT Domain /note= "tetanus toxin helper T cell epitope"
XX FT 38..47
XX FT Domain /note= "LHRH hapten"
XX WO9425060-A1.
XX PD 10-NOV-1994.
XX PF 28-APR-1994; 94WO-US004832.
XX PR 27-APR-1993; 93US-00057166.
XX PR 14-APR-1994; 94US-00229275.
XX PA (LADD/) LADD A E.
XX PA (WANG/) WANG C Y.
XX PA (ZAMB/) ZAMB T.
XX PI Ladd AE, Wang CY, Zamb T;
XX PI MPI, 1994-357910/44.
XX PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
XX PT suppress LHRH activity in males and females.
XX PS Claim 8; Page 88; 213pp; English.
XX CC Synthetic immunogenic peptides are provided in which a universal immune
XX CC stimulator is linked to a peptide or protein hapten containing B cell
XX CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX CC potent immune responses to the coupled peptide or protein. The stimulator
XX CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
XX CC immune response to the coupled peptide in members of a heterogeneous
XX CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
XX CC sequence from the invasin protein of Yersinia. Spacer amino acid
XX CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
XX CC domains and between the immune stimulator and hapten components. When the
XX CC hapten is LHRH, then optionally the invasin domain can be omitted from
XX CC the immune stimulator component. The present sequence represents an LHRH-
XX CC containing immunogenic peptide as above which can be used as a potent
XX CC vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
XX CC carcinoma, prostatic carcinoma, testicular carcinoma, androgen-dependent
XX CC benign uterine tumours, recurrent functional ovarian cysts, (severe)
XX CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX CC induction of infertility. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 47 AA;
XX Query Match 52.8%; Score 76; DB 2; Length 47;
XX Best Local Similarity 88.2%; Pred. No. 8.2e-05;
XX Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RQYIKANSKFTIGTERG 18
Db :|||||
20 KQYIKANSKFTIGTELG 36

```

```

RESULT 12
AAV92638
ID AAV92638 standard; protein; 750 AA.
XX
XX AAV92638;
XX AC
XX DT 10-AUG-2000 (first entry)
XX
XX

```

XX	Mutant human prostate specific membrane antigen construct, hPSM3.1.
DE	
XX	
KW	Prostate specific membrane antigen; immunogenized construct; mutant;
XW	vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW	prostate cancer; cell-associated peptide antigen; foreign epitope.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FT	Key
FT	Location/Qualifiers
FT	Peptide
FT	21..41
FT	/label= P30
FT	/note="Foreign epitope"
FT	213..127
FT	/label= P2
FT	/note="foreign epitope"
PX	
PN	WO200020027-A2.
XX	
PD	13-APR-2000.
PF	05-OCT-1999;
XX	99WO-DK000525.
PR	05-OCT-1998;
PR	98DK-00001261.
PR	20-OCT-1998;
XX	98US-010501IP.
PA	(MEBI-) M & E BIOTECH AS.
XX	
PI	Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
P1	Gautam A, Birk P, Karlsson G;
XX	
DR	WPI; 2000-349917/30.
XX	
PT	Inducing immune responses to weakly immunogenic, tumor associated peptide
PT	antigens for the treatment of breast and prostate cancer.
XX	
FX	Example 1; Page; 220pp; English.
XX	
AA	AAV92627-49 are mutant immunogenized human prostate specific membrane
CC	antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC	The immunogenic analogues of PSM can be used in the claimed method as an
CC	autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC	binding regions and cysteine residues involved in disulfide bonds are
CC	preserved in the immunogenized forms. The method is used for inducing
CC	immune responses against weakly immunogenic cell-associated peptide
CC	antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC	human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC	fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC	simultaneous presentation by antigen producing cells (APCs) of the
CC	animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC	group derived from the PA and/or at least 1 B-cell group derived from the
CC	cell-associated PA; and (2) at least 1 first T helper cell group which is
CC	foreign to the animal. Analogues of human PSM, human Her2 and
CC	human/murine FGF8b comprising a substantial part of all known and
CC	predicted CTL and B-cell epitopes of the respective PA and including at
CC	least one foreign T helper epitope are also claimed. The method is used
CC	to treat prostate, prostate/breast or breast cancer when the PA is human
CC	PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC	from the wild type human PSM (AAV92619), which appears on pages 184-187
CC	of the specification
XX	
AQ	Sequence 750 AA;

Query Match	52.8%	Score 76;	DB 3;	Length 750;
Best Local Similarity	93.8%	Pred. No. 0.0021;		
Matches 15;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	GROYIKANSKFIGITE	16
Dp	211	GNQYIKANSKFIGITE	226

Dö 211 GNÖYIKANSKFIGITE 226

RESULT 13
AA92652 ID AAY92652 standard; peptide; 25 AA.
XX
AC AAY92652;
XX
DT 10-AUG-2000 (first entry)
XX
DE PSMep009 - P2 inserted in hPSM insertion position 10.
XX
KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.
OS Synthetic.
XS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 6..20
FT /label= P2
PN MO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
PR 20-OCT-1998; 98US-0105011P.
PA (MEBI-) M & E BIOTECH AS.
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
DR WPI: 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 1; Page 117; 220pp; English.
XX
CC AAY92650-55 are peptides designed which correspond to the P2 and P30
CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
CC amino acids in each end. The flanking amino acids correspond to the
CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g. T
CC cell proliferation assays, but also for ELISA or other in vitro assays.
CC The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, hPSM,
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method
CC comprises effecting simultaneous presentation by antigen producing cells
CC (APCs) of the animals immune system of: (1) at least 1 CTL (cytotoxic T-
CC lymphocyte) group derived from the PA and/or at least 1 B-cell group
CC derived from the cell-associated PA; and (2) at least 1 first T helper
CC cell group which is foreign to the animal. Analogues of human PSM, human
CC Her2 and human/murine FGF8b comprising a substantial part of all known
CC and predicted CTL and B-cell epitopes of the respective PA and including
CC at least one foreign T helper epitope (e.g. P2 and/or P30) are also
CC claimed. The method is used to treat prostate, prostate/breast or breast
CC cancer when the PA is human PSM, FGF8b and Her2, respectively
XX

Sequence 25 AA;

Query Match 52.1%; Score 75; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

2 ROYIKANSKEFIGITE 16
|||||
5 ROYIKANSKEFIGITE 19
|||||

Query Match	52.1%	Score 75;	DB 3;	Length 25;
Best Local Similarity	100.0%;	Pred. No. 5.7e-05;		
Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.1

QY	2 RØYKANSKFICITE 16
Dø	5 RØYKANSKFICITE 19

Db 5 ROYKANSKFIGITE 19

RESULT 14
 ID AAY92651 standard; peptide; 25 AA.
 AC AAY92651;
 XX 10-AUG-2000 (first entry)
 DT
 DE PSM008 - P2 inserted in hPSM insertion position 8.
 XX
 KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
 KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
 KW prostate cancer; cell-associated peptide antigen.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT peptide 6.20
 FT /label= P2
 XX
 WO200020027-A2.
 XX
 PN 13-APR-2000.
 XX
 PF 05-OCT-1999; 99MO-DK000525.
 XX
 PR 05-OCT-1998; 98DK-00001261.
 PR 20-OCT-1998; 98US-0105011P.
 XX
 XX (MEBI-) M & E BIOTECH AS.
 PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 DR WPI; 2000-349917/30.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated peptide
 PT antigens for the treatment of breast and prostate cancer.
 XX
 PS Example 1; Page 117; 220pp; English.
 XX
 CC AAY92650-55 are peptides designed which correspond to the P2 and P30
 CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
 CC amino acids in each end. The flanking amino acids correspond to the
 CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g. T
 CC cell proliferation assays, but also for ELISA or other in vitro assays.
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, hPSM,
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The method
 CC comprises effecting simultaneous presentation by antigen producing cells
 CC (APCs) of the animals immune system of: (1) at least 1 CTL (cytotoxic T-
 CC lymphocyte) group derived from the PA and/or at least 1 B-cell group
 CC derived from the cell-associated PA; and (2) at least 1 first T helper
 CC cell group which is foreign to the animal. Analogues of human PSM, human
 CC Her2 and human/murine FGF8b comprising a substantial part of all known
 CC and predicted CTL and B-cell epitopes of the respective PA and including
 CC at least one foreign T helper epitope (e.g. P2 and/or P30) are also
 CC claimed. The method is used to treat prostate, prostate/breast or breast
 CC cancer when the PA is human PSM, FGF8b and Her2, respectively
 XX
 SQ Sequence 25 AA:
 Query Match 52.1%; Score 75; DB 3; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RYIKANSKFIGITTE 16
 DB 5 RYIKANSKFIGITTE 19

RESULT 15
 ID ABB79179 standard; peptide; 31 AA.
 AC ABB79179;
 XX 07-AUG-2002 (first entry)
 DT
 DE Human cytomegalovirus PP150 related vaccine peptide SEQ ID NO:5.
 XX
 KW Human cytomegalovirus PP150; HCMV; human leukocyte antigen; HLA; CTL;
 KW cytotoxic T lymphocyte; immunoreactive; epitope; vaccine; infection;
 KW immunogenic; virucide; immune response.
 XX
 OS Human cytomegalovirus.
 OS
 PN WO200234769-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 22-OCT-2001; 2001WO-US032589.
 XX
 PR 20-OCT-2000; 2000US-0241944P.
 XX
 PA (CITY) CITY OF HOPE.
 XX
 PI Diamond DJ;
 XX
 DR WPI; 2002-471432/50.
 XX
 PT New immunoreactive peptide cytotoxic T-lymphocytes epitope of human
 PT cytomegalovirus PP150, useful for creating vaccines against
 PT cytomegalovirus infection.
 XX
 PS Disclosure; Page 14; 28pp; English.
 XX
 CC The present invention describes a peptide (I) which is an immunogenic
 CC epitope recognised by CD8+ class I major histocompatibility complex (MHC)
 CC restricted cytotoxic T-lymphocytes (CTL) of patients harbouring latent
 CC cytomegalovirus (HCMV) infection. Also described are: (1) a vaccine (II)
 CC against human cytomegalovirus (CMV), which comprises (i); (2) a cellular
 CC vaccine (III) against human CMV comprising antigen presenting cells which
 CC present (i); (3) a recombinant viral vector (IV) which contains a gene
 CC encoding (i); and (4) an immunological reagent (V) which comprises (i).
 CC (i) has virucide activity. (ii) or (iii) are useful for modulating the
 CC immune response to HCMV infection, and for vaccinating a mammal in need
 CC of it against HCMV. (i) is useful for eliciting a cellular immune against
 CC HCMV by normal and immunodeficient subjects, for creating efficient
 CC vaccines against HCMV, and in immunological methods to detect PP150-
 CC reactive CTL in a patient or a sample from the patient. The present
 CC sequence represents a HCMV PP150 related vaccine peptide, which is given
 CC in the exemplification of the present invention
 XX
 SQ Sequence 31 AA:
 Query Match 52.1%; Score 75; DB 5; Length 31;
 Best Local Similarity 63.0%; Pred. No. 7.4e-05;
 Matches 17; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
 QY 3 QYIKANSKFIGITTEG-RILKEPVHG 27
 DB 4 QYIKANSKFIGITEAAAGTWTSTPVQG 30

Search completed: July 20, 2004, 06:30:27
 Job time : 42.7705 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 11.0164 Seconds

(without alignments)
279.413 Million cell updates/sec

Title: US-09-673-166B-275

Sequence: 168

Sequence: 1 GROYTKANSKPIGITERGRNPDIVYQYMDL 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	78	46.4	210	2 F47330	RNA-directed DNA p
2	75	44.6	210	2 D47330	RNA-directed DNA p
3	75	44.6	210	2 A47330	RNA-directed DNA p
4	75	44.6	210	2 B47330	RNA-directed DNA p
5	75	44.6	210	2 C47330	RNA-directed DNA p
6	75	44.6	210	2 E47330	RNA-directed DNA p
7	75	44.6	219	2 S32139	RNA-directed DNA p
8	75	44.6	219	2 S32071	RNA-directed DNA p
9	75	44.6	219	2 S32079	RNA-directed DNA p
10	75	44.6	219	2 S32077	RNA-directed DNA p
11	75	44.6	219	2 S32075	RNA-directed DNA p
12	75	44.6	219	2 S32074	RNA-directed DNA p
13	75	44.6	219	2 S32073	RNA-directed DNA p
14	75	44.6	219	2 S32070	RNA-directed DNA p
15	75	44.6	219	2 S32080	RNA-directed DNA p
16	75	44.6	219	2 S32072	RNA-directed DNA p
17	75	44.6	219	2 S32134	RNA-directed DNA p
18	75	44.6	219	2 S32138	RNA-directed DNA p
19	75	44.6	219	2 S32122	RNA-directed DNA p
20	75	44.6	219	2 S32117	RNA-directed DNA p
21	75	44.6	219	2 S32120	RNA-directed DNA p
22	75	44.6	219	2 S32096	RNA-directed DNA p
23	75	44.6	219	2 S32066	RNA-directed DNA p
24	75	44.6	219	2 S32140	RNA-directed DNA p
25	75	44.6	219	2 S32119	RNA-directed DNA p
26	75	44.6	219	2 S32063	RNA-directed DNA p
27	75	44.6	219	2 S32137	RNA-directed DNA p
28	75	44.6	219	2 S32059	RNA-directed DNA p
29	75	44.6	219	2 S32053	RNA-directed DNA p

30	75	44.6	219	2 S32095	RNA-directed DNA p
31	75	44.6	219	2 S32056	RNA-directed DNA p
32	75	44.6	219	2 S32064	RNA-directed DNA p
33	75	44.6	219	2 S32054	RNA-directed DNA p
34	75	44.6	219	2 S32062	RNA-directed DNA p
35	75	44.6	219	2 S32061	RNA-directed DNA p
36	75	44.6	219	2 S32058	RNA-directed DNA p
37	75	44.6	219	2 S32060	RNA-directed DNA p
38	75	44.6	219	2 S32057	RNA-directed DNA p
39	75	44.6	219	2 S32159	RNA-directed DNA p
40	75	44.6	219	2 S32157	RNA-directed DNA p
41	75	44.6	219	2 S32160	RNA-directed DNA p
42	75	44.6	219	2 S32078	RNA-directed DNA p
43	75	44.6	219	2 S32135	RNA-directed DNA p
44	75	44.6	219	2 S32069	RNA-directed DNA p
45	75	44.6	219	2 S32126	RNA-directed DNA p

ALIGNMENTS

RESULT 1

F47330
RNA-directed DNA polymerase (BC 2.7.7.49) - human immunodeficiency virus type 1 (strain N/Alternate names: DNA nucleotidyltransferase (RNA-directed) ; reverse transcriptase; rev C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C/Accession: F47330
R/Shirasaka, T.; Yarchuan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.; Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A/Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during therapy
A/Reference number: A47330; MUID:93133828; PMID:8380641
A/Accession: F47330
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-210 <SH1>
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase

Query Match
Best Local Similarity 69.2%; Score 78; DB 2; Length 210;
Matches 18; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 9 SKFIGITE--RGRNPDIVYQYMDL 32
| | | | | : | | | | | | | | | |
Db 142 SSMIKLEPRKQNPDIYQYMDL 167

RESULT 2

D47330
RNA-directed DNA polymerase (BC 2.7.7.49) - human immunodeficiency virus type 1 (strain N/Alternate names: DNA nucleotidyltransferase (RNA-directed) ; reverse transcriptase; rev C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C/Accession: D47330
R/Shirasaka, T.; Yarchuan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.; Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A/Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during therapy
A/Reference number: A47330; MUID:93133828; PMID:8380641
A/Accession: D47330
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-210 <SH1>
C/Superfamily: pol polyprotein
C/Keywords: nucleotidyltransferase; polymorphism

Query Match
Best Local Similarity 44.6%; Score 75; DB 2; Length 210;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 RGRNPDIVYQYMDL 32
| : | | | | | | | | | | | | | | | |
Db 152 RKQNPDIYQYMDL 167

```
RESULT 3
A47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N:Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C:Accession: A47330
R:Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A:Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during thera
A:Reference number: A47330; PMID:93133828; PMID:8380641
A:Accession: A47330
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-210 <SHI>
C:Superfamily: pol polypeptide
C:Keywords: nucleotidyltransferase

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RQNPDIIVYQYMDL 167

RESULT 4
B47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N:Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C:Accession: B47330
R:Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A:Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during thera
A:Reference number: A47330; PMID:93133828; PMID:8380641
A:Accession: B47330
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-210 <SHI>
C:Superfamily: pol polypeptide
C:Keywords: nucleotidyltransferase; polymorphism

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RQNPDIIVYQYMDL 167

RESULT 5
C47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N:Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C:Accession: C47330
R:Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A:Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during thera
A:Reference number: A47330; PMID:93133828; PMID:8380641
A:Accession: C47330
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-210 <SHI>
C:Superfamily: pol polypeptide
C:Keywords: nucleotidyltransferase; polymorphism

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RQNPDIIVYQYMDL 167
```

```
Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RQNPDIIVYQYMDL 167

RESULT 6
B47330
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (strain
N:Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 22-Apr-1995
C:Accession: B47330
R:Shirasaka, T.; Yarchan, R.; O'Brien, M.C.; Husson, R.N.; Anderson, B.D.; Kojima, E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 562-566, 1993
A:Title: Changes in drug sensitivity of human immunodeficiency virus type 1 during ther
A:Reference number: A47330; PMID:93133828; PMID:8380641
A:Accession: B47330
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-210 <SHI>
C:Superfamily: pol polypeptide
C:Keywords: nucleotidyltransferase; polymorphism

Query Match          44.6%; Score 75; DB 2; Length 210;
Best Local Similarity 87.5%; Pred. No. 0.0045;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 152 RQNPDIIVYQYMDL 167

RESULT 7
S32139
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32139
R:Wildemann, B.; Haas, J.; Hahn, M.; Ehrhart, K.; Storch-Hagenlocher, B.
submitted to the EMBL Data Library, February 1993
A:Description: In vivo occurrence of drug resistance mutations under prolonged zidovudine
A:Reference number: S32117
A:Accession: S32139
A:Molecule type: DNA
A:Residues: 1-219 <MTL>
A:Cross-references: EMBL:X71106; NID:G287982; PIDN:CA50423.1; PID:G287983
C:Superfamily: pol polypeptide
C:Keywords: nucleotidyltransferase

Query Match          44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
|:|||||
Db 172 RQNPDIIVYQYMDL 187

RESULT 8
S32071
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32071
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
```

A:Reference number: S32047
A:Accession: S32071
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70606; NID:g287694; PIDN:CAA49960.1; PID:g938201
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
|:|||||
Db 172 RKQNPDIYQYMDL 187

RESULT 9
S32079
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32079
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32079
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70610; NID:g287702; PIDN:CAA49964.1; PID:g938209
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
|:|||||
Db 172 RKQNPDIYQYMDL 187

RESULT 10
S32077
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32077
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32077
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70609; NID:g287700; PIDN:CAA49963.1; PID:g938207
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
|:|||||
Db 172 RKQNPDIYQYMDL 187

RESULT 11
S32075

RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32075
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32075
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70608; NID:g287698; PIDN:CAA49962.1; PID:g938205
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
|:|||||
Db 172 RKQNPDIYQYMDL 187

RESULT 12
S32074
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32074
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32074
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70600; NID:g287697; PIDN:CAA49954.1; PID:g938204
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
|:|||||
Db 172 RKQNPDIYQYMDL 187

RESULT 13
S32073
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S32073
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32073
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70607; NID:g287696; PIDN:CAA49961.1; PID:g938203
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIVYQYMDL 32
| : |||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 14

S32070
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
Submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32070
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70598; NID:g287693; PIDN:CAA49952.1; PID:g938200
A:Accession: S32067
A:Molecule type: DNA
A:Residues: 1-219 <WIL2>
A:Cross-references: EMBL:X70596; NID:g287689; PIDN:CAA49950.1; PID:g938197
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIVYQYMDL 32
| : |||||
Db 172 RKQNPDIYIYQYMDL 187

RESULT 15

S32080
RNA-directed DNA polymerase (EC 2.7.7.49) - human immunodeficiency virus type 1 (fragment)
N:Alternate names: reverse transcriptase
C:Species: human immunodeficiency virus type 1, HIV-1
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
R:Wildemann, B.; Haas, J.; Ehrhart, K.; Hahn, M.; Storch-Hagen, B.
Submitted to the EMBL Data Library, January 1993
A:Description: In vivo comparison of zidovudine resistance in blood and CSF of HIV-1 inf
A:Reference number: S32047
A:Accession: S32080
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:X70611; NID:g287703; PIDN:CAA49965.1; PID:g938210
C:Superfamily: pol polyprotein
C:Keywords: nucleotidyltransferase

Query Match 44.6%; Score 75; DB 2; Length 219;
Best Local Similarity 87.5%; Pred. No. 0.0047;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIVYQYMDL 32
| : |||||
Db 172 RKQNPDIYIYQYMDL 187

Search completed: July 20, 2004, 06:34:10
Job time : 11.0164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 6.29508 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166B-275

Sequence: 1 GRQYIKANSKFGITGRGNPDIVITYQWDDL 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141661 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match	Length	DB ID	Description
1	75	44.6	1003	1	P03369 human immun
2	75	44.6	1003	1	P04585 human immun
3	75	44.6	1003	1	P12497 human immun
4	75	44.6	1003	1	P20892 human immun
5	75	44.6	1006	1	P05961 human immun
6	75	44.6	1015	1	P04587 human immun
7	75	44.6	1015	1	P03367 human immun
8	75	44.6	1015	1	P03368 human immun
9	74	43.5	1007	1	P20875 human immun
10	74	43.5	1003	1	P35563 human immun
11	72	42.9	1002	1	P04588 human immun
12	72	42.9	1002	1	P24740 human immun
13	72	42.9	1015	1	P03366 human immun
14	72	42.9	1027	1	P17283 chimpanzee
15	71	42.3	1002	1	P18802 human immun
16	71	42.3	1002	1	P12499 human immun
17	70	41.7	1314	1	P04598 clostridium
18	68	40.5	1002	1	P05959 human immun
19	68	40.5	1002	1	P12451 human immun
20	61	36.3	1035	1	P20876 human immun
21	60	35.7	1055	1	P15561 bovine immu
22	60	35.7	1056	1	P19560 bovine immu
23	60	35.7	1056	1	P15561 bovine immu
24	59	35.1	1035	1	P18042 human immun
25	59	35.1	1049	1	P17757 human immun
26	59	35.1	1073	1	P18096 human immun
27	59	35.1	1142	1	P11204 equine infe
28	58	34.5	1145	1	P12502 simian immu
29	58	34.5	1146	1	P15505 simian immu
30	58	34.5	1146	1	P05897 simian immu
31	56	33.3	1019	1	
32	56	33.3	1022	1	
33	56	33.3	1054	1	

34	56	33.3	1056	1	POL_SIVM1	P05896 simian immu
35	55	32.7	1124	1	POL_FIVPE	P16088 feline immu
36	55	32.7	1124	1	POL_FIVSD	P19028 feline immu
37	53	31.5	1009	1	POL_SIVGB	P22382 simian immu
38	53	31.5	1034	1	POL_SIVBZ	P24107 human immun
39	53	31.5	1035	1	POL_HV2CA	P05962 human immun
40	53	31.5	1036	1	POL_HV2RO	P04584 human immun
41	53	31.5	1058	1	POL_HV2D2	P15833 human immun
42	53	31.5	1165	1	POL_GALV	P21414 gibbon ape
43	52	31.0	481	1	GSHI_CIOAB	O971V1 clostridium
44	51	30.4	206	1	URK_LACIA	O9CF21 lactococcus
45	51	30.4	1196	1	POL_MIVRD	P11227 radiation m

ALIGNMENTS

RESULT 1
ID POL_HV1A2 STANDARD; PRT; 1003 AA.
AC P03369;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropesin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)];
DE POL.
OS Human immunodeficiency virus type 1 (ARV2/SP2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85090453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
RT (ARV-2).";
RL Science 227:484-492(1985).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 57-155.
RX MEDLINE=89346747; PubMed=2548279;
RA Wlodawer A., Miller M., Jaskolski M., Suthanarayana B.K.,
RA Baldwin E., Weber I.T., Selk L.M., Clawson L., Schneider J.,
RA Kent S.B.H.;
RT "Conserved folding in retroviral proteases: crystal structure of a
RT synthetic HIV-1 protease.";
RL Science 245:616-621(1998).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS) OF 57-155 OF COMPLEX WITH INH.
RA Abbenante G., March D.R., Bergman D.A., Hunt P.A.,
RA Garham B., Dancer R.J., Martin J.L., Fairlie D.P.,
RT "Regioselective structural and functional mimicry of peptides - design
RT of hydrolytically-stable cyclic peptidomimetic inhibitors of HIV-1
RT protease.";
RL J. Am. Chem. Soc. 117:10220-10226(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS) OF 57-155 OF COMPLEX WITH INH.
RA March D.R., Abbenante G., Bergman D.A., Brinkworth R.I.,
RA Wickramasinghe W., Begun J., Martin J.L., Fairlie D.P.;
RT "Substrate-based cyclic peptidomimetics of Phe-1le-Val that inhibit
RT HIV-1 protease using a novel enzyme-binding mode.";
RL J. Am. Chem. Soc. 118:3375-3379(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 57-155.
RX MEDLINE=96438794; PubMed=8841139;
RA Rose R.B., Craik C.S., Douglas N.L., Stroud R.M.;
RT "Three-dimensional structures of HIV-1 and SIV protease product
RT complexes.";
RL Biochemistry 35:12933-12944(1996).
CC -I- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -I- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5' -

phosphomonoester.
 -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + {DNA}(N).
 -1- PTM: Cleavage sites that yield the mature proteins remain to be determined.
 -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).

 EMBL: K02007; AAB59876.1; -
 PIR: A03968; GNWMA2.
 PDB: 3HYF; 15-JAN-90.
 PDB: 1CPI; 08-MAR-96.
 PDB: 1MTR; 01-AUG-96.
 PDB: 1YTG; 12-MAR-97.
 PDB: 1YTH; 12-MAR-97.
 PDB: 1AID; 15-OCT-97.
 PDB: 1B6F; 07-JAN-00.
 PDB: 1B6K; 07-JAN-00.
 PDB: 1B6L; 07-JAN-00.
 PDB: 1B6M; 07-JAN-00.
 PDB: 1B6N; 07-JAN-00.
 PDB: 1B6O; 07-JAN-00.
 PDB: 1B6P; 07-JAN-00.
 PDB: 1B4K; 11-OCT-00.
 PDB: 1D4L; 01-NOV-00.
 PDB: 1F7A; 27-JUN-01.
 PDB: 1KJ4; 06-MAR-02.
 PDB: 1KJ7; 06-MAR-02.
 PDB: 1KJP; 06-MAR-02.
 PDB: 1KJG; 06-MAR-02.
 PDB: 1KJH; 06-MAR-02.
 PDB: 1KZK; 03-APR-02.
 PDB: 1MT7; 07-JAN-03.
 PDB: 1MT8; 07-JAN-03.
 PDB: 1MT9; 07-JAN-03.
 PDB: 1MT3; 07-JAN-03.
 PDB: 1N49; 07-JAN-03.
 PDB: 2AID; 15-OCT-97.
 PDB: 3AID; 17-SEP-97.
 PDB: 4HYF; 15-OCT-92.
 PDB: 8HYF; 31-OCT-93.
 HIV: K02007; POLSF2.
 MEROPS: A02.001; -
 InterPro: IPR001969; Asparticase_AS.
 InterPro: IPR001037; Integrase_C.
 InterPro: IPR003308; Integrase_Zn.
 InterPro: IPR009007; Pept_A acid.
 InterPro: IPR001995; Peptidase_A2.
 InterPro: IPR002156; RNaseH.
 InterPro: IPR001584; Rve.
 InterPro: IPR004777; RVTse.
 Pfam: PF00552; Integrase; 1.
 Pfam: PF02022; Integrase; 1.
 Pfam: PF00075; rnaseH; 1.
 Pfam: PF00665; rve; 1.
 Pfam: PF00077; tvp; 1.
 Pfam: PF00078; tvp; 1.
 PROSITE: PS00141; ASP_PROTEASE; 1.
 PROSITE: PS00175; ASP_PROT_RETROV; 1.
 AIDS: Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease; Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
 CHAIN 57 155
 ACT SITE 81 155
 STRAND 58 59
 STRAND 66 71
 BY SIMILARITY.

FT TURN 72 73
 FT STRAND 74 80
 FT TURN 82 83
 FT STRAND 88 89
 FT STRAND 96 93
 FT TURN 99 105
 FT STRAND 106 107
 FT STRAND 108 122
 FT TURN 123 124
 FT STRAND 125 133
 FT STRAND 125 133
 FT STRAND 139 141
 FT HELIX 143 146
 FT TURN 147 150
 FT STRAND 152 154
 SQ SEQUENCE 1003 AA; 113723 MW; 1519A67242219CB3 CRC64;
 Query Match 44.6%; Score 75; DB 1; Length 1003;
 Best Local Similarity 87.5%; Pred. No. 0.0037;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 17 RGRNPDIYVOYMDL 32
 Db 327 RKQNPDIYVOYMDL 342
 ID POL_HV1H2 STANDARD; PRT; 1003 AA.
 AC P04585; O09777; Q9WUC5;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE POL polyprotein [contains: Protease (Retriopepsin) (EC 3.4.23.16); Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN POL.
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 NC NCBI_Taxid=11706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6729196; PubMed=3040055;
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S., Gallo R.C., Wong-Staal F.;
 RT "Complete nucleotide sequences of functional clones of the AIDS virus";
 RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
 RN [2]
 RP REVISIONS.
 RA Ogata N., Alter H.J., Miller R.H., Purcell R.H.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chappay C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
 RX MEDLINE=99043699; PubMed=9827997;
 RA Kervinen J., Lubkowski J., Zdanov A., Bharti D., Dunn B.M., Hui K.Y., Powell D.J., Kay J., Mlodner A., Gustchina A.;
 RT "Toward a universal inhibitor of retroviral proteases: comparative analysis of the interactions of LP-130 complexed with proteases from HIV-1, FIV, and EIAV";
 RL Protein Sci. 7:2314-2323(1998).
 RN [5]
 RP STRUCTURE BY NMR OF 57-155.
 RX MEDLINE=9702126; PubMed=868486;
 RA Yamazaki T., Hinck A.P., Wang Y.-X., Nicholson L.K., Torchia D.A., Lam P.Y.S.;
 RT "Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy";
 RL Protein Sci. 5:495-506(1996).
 RN [6]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 156-595.
RX MEDLINE=96097398; PubMed=8535785;
RA Ren J., Ennouf R.M., Hopkins A.L., Ross C., Jones E.Y., Stammers D.K.,
RA Stuart D.I.;
RT "The structure of HIV-1 reverse transcriptase complexed with
RT 9-chloro-TIBO: lessons for inhibitor design."; <http://www.isd-sib.ch/announce/>
RL Structure 3:915-926(1995).
[7]
RP X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS) OF 156-595.
RX MEDLINE=96208551; PubMed=8648598;
RA Hopkins A.L., Ren J., Ennouf R.M., Willcox B.E., Jones E.Y., Ross C.,
RA Miyasaka T., Walker R.T., Tanaka H., Stammers D.K., Stuart D.I.;
RT "Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT
RT series reveal conformational changes relevant to the design of potent
RT non-nucleoside inhibitors";
RL J. Med. Chem. 39:1589-1600(1996).
[8]
RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 157-694.
RX MEDLINE=97268683; PubMed=9108091;
RA Ennouf R.M., Ren J., Hopkins A.L., Ross C.K., Jones E.Y.,
RA Stammers D.K., Stuart D.I.;
RT "Unique features in the structure of the complex between HIV-1
RT reverse transcriptase and the bis(heteroaryl)piperazine (BHAP)
RT U-90152 explain resistance mutations for this nonnucleoside
RT inhibitor";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3984-3989(1997).
[9]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 157-694.
RX MEDLINE=98356189; PubMed=9689112;
RA Ren J., Ennouf R.M., Hopkins A.L., Jones E.Y., Kirby I., Keeling J.,
RA Ross C.K., Larder B.A., Stuart D.I., Stammers D.K.;
RT "3'-Azido-3'-deoxythymidine drug resistance mutations in HIV-1
RT reverse transcriptase can induce long range conformational changes."; <http://www.isd-sib.ch/announce/>
RL Proc. Natl. Acad. Sci. U.S.A. 95:9518-9523(1998).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isd-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; K03455; AAB50259.1; ALT_INIT.
DR EMBL; AF033819; AAC82598.2; -.
DR PDB; 1REV; 14-OCT-96.
DR PDB; 1RT1; 21-APR-97.
DR PDB; 1RT2; 21-APR-97.
DR PDB; 1RT3; 16-FEB-99.
DR PDB; 1RTH; 03-APR-96.
DR PDB; 1RTJ; 03-APR-96.
DR PDB; 1RTJ; 03-APR-96.
DR PDB; 1BEV; 17-AUG-96.
DR PDB; 1BVG; 17-AUG-96.
DR PDB; 1KLM; 18-MAR-98.
DR PDB; 1A30; 29-APR-98.
DR PDB; 1ODY; 16-FEB-99.
DR PDB; 1BV7; 14-JAN-00.
DR PDB; 1BV9; 12-JAN-00.
DR PDB; 1BWA; 12-JAN-00.
DR PDB; 1BWB; 12-JAN-00.
DR PDB; 1COT; 19-JUL-00.
DR PDB; 1COW; 19-JUL-00.

DR PDB; 1CIB; 21-JUL-00.
DR PDB; 1CIC; 06-SEP-00.
DR PDB; 1DMP; 12-NOV-97.
DR PDB; 1DTQ; 20-MAR-00.
DR PDB; 1DTT; 02-APR-00.
DR PDB; 1EP4; 27-SEP-00.
DR PDB; 1EX4; 26-JUL-00.
DR PDB; 1EXQ; 03-NOV-00.
DR PDB; 1FK9; 28-FEB-01.
DR PDB; 1FKO; 28-FEB-01.
DR PDB; 1FKP; 03-NOV-00.
DR PDB; 1HVP; 06-APR-99.
DR PDB; 1HYR; 15-MAY-95.
DR PDB; 1HWR; 23-MAR-99.
DR PDB; 1HXB; 12-MAR-97.
DR PDB; 1JXH; 03-OCT-01.
DR PDB; 1JUA; 03-OCT-01.
DR PDB; 1JUB; 03-OCT-01.
DR PDB; 1JUC; 03-OCT-01.
DR PDB; 1JUF; 03-OCT-01.
DR PDB; 1JUG; 03-OCT-01.
DR PDB; 1JUG; 22-AUG-01.
DR PDB; 1JW0; 30-OCT-02.
DR PDB; 1JW2; 30-OCT-02.
DR PDB; 1JWC; 30-OCT-02.
DR PDB; 1JWE; 30-OCT-02.
DR PDB; 1JWF; 30-OCT-02.
DR PDB; 1JER; 15-APR-98.
DR PDB; 1JES; 15-APR-98.
DR PDB; 1MET; 15-APR-98.
DR PDB; 1MEU; 15-APR-98.
DR PDB; 1OIW; 18-FEB-03.
DR PDB; 1ODW; 01-APR-97.
DR PDB; 1OBR; 15-OCT-97.
DR PDB; 1OBS; 15-OCT-97.
DR PDB; 1OBT; 15-OCT-97.
DR PDB; 1OBU; 15-OCT-97.
DR PDB; 1RT4; 29-JUL-99.
DR PDB; 1RT5; 29-JUL-99.
DR PDB; 1RT6; 29-JUL-99.
DR PDB; 1RT7; 29-JUL-99.
DR PDB; 1RTD; 12-JAN-00.
DR PDB; 1VRT; 03-APR-96.
DR PDB; 3PHV; 15-JAN-92.
DR HIV; K03455; POLSHXR2.
DR MEROPS; A02.001.-;
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR00308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR00477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvc; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS01075; ASP_PROT_RETROV; 1.
KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 57 155
FT ACT_SITE 81 81
FT STRAND 58 59
FT STRAND 66 71
FT TURN 72 73
FT STRAND 74 80
FT TURN 82 83
FT STRAND 88 89
FT STRAND 99 104

```

FT STRAND 109 122
FT STRAND 126 133
FT STRAND 140 141
FT HELIX 143 146
FT TURN 147 150
FT STRAND 152 154
FT STRAND 167 167

Query Match 44.6%; Score 75; DB 1; Length 1003;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGNPDIVITYQYMDL 32
Db 327 RKQNPDIIVYQYMDL 342

RESULT 3
POL_HV1NS STANDARD; PRT; 1003 AA.
ID POL_HV1NS
AC P12497;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polypeptide [contains: Protease (retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
POL.
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11698;
RN [1]
RP SEQUENCE FROM N.A. (CLONE PNL4-3).
RA Buckler C.E., Buckler-White A.U., Willey R.L., McCoy J.;
Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.
RL [2]
RX MEDLINE=90354401; PubMed=2201682;
RA Fitzgerald P.M.D., McKeever B.M., van Middlesworth J.F.,
Springer J.P., Heimbach J.C., Leu C.-T., Herber W.K., Dixon R.A.F.,
Darke P.L.;
"Crystallographic analysis of a complex between human
immunodeficiency virus type 1 protease and acetyl-peptstatin at 2.0-A
resolution.";
RT J. Biol. Chem. 265:14209-14219 (1990).
RL U. Biol. Chem. 265:14209-14219 (1990).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19921; AAA44988.1; -.
DR PDB; 5HP; 15-OCT-91.
DR PDB; 4PHV; 31-OCT-93.
DR PDB; 1B9F; 19-JUL-99.
DR PDB; 1BHT; 28-OCT-98.
DR PDB; 1B14; 18-NOV-98.
DR PDB; 1B15; 16-SEP-98.
DR PDB; 1B1U; 16-SEP-98.
DR PDB; 1WJB; 13-MAY-98.
DR PDB; 1WJD; 13-MAY-98.

```

```

DR PDB; 21TG; 12-MAR-97.
DR PDB; 9HP; 15-JUL-92.
DR HIV; M19921; POLSML43.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Peptidase_Zn.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR004477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00075; RNaseH_1.
DR Pfam; PF00665; Rve_1.
DR Pfam; PF00077; RVD_1.
DR Pfam; PF00078; RVT_1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW AIDS; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 57 155
FT ACT_SITE 81 81
FT STRAND 58 60
FT STRAND 66 71
FT TURN 72 73
FT STRAND 74 80
FT TURN 82 83
FT STRAND 87 90
FT STRAND 99 105
FT TURN 106 107
FT STRAND 108 122
FT TURN 123 124
FT STRAND 125 134
FT STRAND 140 141
FT HELIX 143 146
FT HELIX 147 149
FT TURN 150 150
FT STRAND 152 154
SQ SEQUENCE 1003 AA; 113535 MW; EB9CDBA30E114A55 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1003;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGNPDIVITYQYMDL 32
Db 327 RKQNPDIIVYQYMDL 342

RESULT 4
POL_HV10Y STANDARD; PRT; 1003 AA.
ID POL_HV10Y
AC P20892;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polypeptide [contains: Protease (retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (OVI isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11698;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90148544; PubMed=2559749;
Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Main-Hobson S.;
"A highly defective HIV-1 strain isolated from a healthy Gabonese
individual presenting an atypical western blot.";
RT AIDS 3:707-715 (1989).
RL AIDS 3:707-715 (1989).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.

```



```
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonester
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC HEALTHY GABONESE INDIVIDUAL.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26727; AAA83392.1; -.
CC HSSP: P03366; 1RVL.
CC HIV: M26727; POL5OYL.
CC MEROPS: A02.001; -.
CC InterPro: IPR001969; Aspprotease_AS.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR003308; Integrase_Zn.
CC InterPro: IPR009007; Pept_A_acid.
CC InterPro: IPR001995; Peptidase_A2.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR001584; Rve.
CC InterPro: IPR000477; RVTse.
CC Pfam: PF00552; Integrase_1.
CC Pfam: PF02022; Integrase_Zn; 1.
CC Pfam: PF00075; rnaseh; 1.
CC Pfam: PF00665; rve; 1.
CC Pfam: PF00077; rvp; 1.
CC Pfam: PF00078; rvt; 1.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
CC PROSITE: PS50175; ASP_PROT_RETROV; 1.
CC AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC CHAIN 57 155 BY SIMILARITY.
CC FT ACT SITE 81 81 IN-PRIME TERMINATION CODON.
CC SQ SEQUENCE 1003 AA; 113718 MW; AFB997A0ED88A98 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1003;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
Db 327 RKQNPDIYIYQYMDL 342

RESULT 5
POL_HVIMN STANDARD; PRT; 1006 AA.
AC P05961;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [Contains: Protease (Retriopepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]].
GN POL.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCB1_TaxID=11696;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88219542; PubMed=3369091; Aldovini A., Collalti E.,
RA Guirgo C., Guo H.-G., Franchini G., Gallo R.C., Reltz K., Jr.;
RA Farrell K., Wong-Staal F., Gallo R.C., Reltz K., Jr.;
RA "Envelope sequences of two new United States HIV-1 isolates.";
```

```
RL Virology 164:531-536(1988).
CC -1- CATALYTIC ACTIVITY: Specific for a PL residue that is hydrophobic,
CC and PL' variable, but often pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17449; -, NOT ANNOTATED_CDS.
CC HSSP: P03366; 1RVL.
CC HIV: M17449; POL5MN.
CC MEROPS: A02.001; -.
CC InterPro: IPR001969; Aspprotease_AS.
CC InterPro: IPR001037; Integrase_C.
CC InterPro: IPR003308; Integrase_Zn.
CC InterPro: IPR009007; Pept_A_acid.
CC InterPro: IPR001995; Peptidase_A2.
CC InterPro: IPR002156; RNaseH.
CC InterPro: IPR001584; Rve.
CC InterPro: IPR000477; RVTse.
CC Pfam: PF00552; Integrase_1.
CC Pfam: PF02022; Integrase_Zn; 1.
CC Pfam: PF00075; rnaseh; 1.
CC Pfam: PF00665; rve; 1.
CC Pfam: PF00077; rvp; 1.
CC Pfam: PF00078; rvt; 1.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
CC PROSITE: PS50175; ASP_PROT_RETROV; 1.
CC AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC CHAIN 60 158 BY SIMILARITY.
CC FT ACT SITE 84 84 IN-PRIME TERMINATION CODON.
CC SQ SEQUENCE 1006 AA; 113860 MW; 460C50DF92AF9B3 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1006;
Best Local Similarity 87.5%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYIYQYMDL 32
Db 330 RKQNPDIYIYQYMDL 345

RESULT 6
POL_HV1B5 STANDARD; PRT; 1015 AA.
AC P04587;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retriopepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]].
GN POL.
OS Human immunodeficiency virus type 1 (BHS isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCB1_TaxID=11682;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Ratajski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Pappas T.S., Chirayeb J., Chang N.T., Gallo R.C.,
RA Wons-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02012; AAA44653.1; -.
DR PDB; 1BDL; 12-AUG-98.
DR PDB; 1BDO; 12-AUG-98.
DR PDB; 1BDR; 14-OCT-98.
DR PDB; 1FED; 01-JUN-01.
DR PDB; 1FEO; 01-JUN-01.
DR PDB; 1FEF; 01-JUN-01.
DR PDB; 1FFI; 01-JUN-01.
DR PDB; 1FG6; 01-JUN-01.
DR PDB; 1FG8; 01-JUN-01.
DR PDB; 1FGC; 01-JUN-01.
DR PDB; 1G2K; 16-NOV-01.
DR PDB; 1HVV; 31-MAR-95.
DR PDB; 1HVF; 31-JUL-94.
DR PDB; 1HVJ; 30-APR-94.
DR PDB; 1HVP; 15-JAN-95.
DR PDB; 1HVS; 14-FEB-95.
DR PDB; 1HVT; 10-JUL-02.
DR PDB; 1KIT; 10-JUL-02.
DR PDB; 1K2B; 10-JUL-02.
DR PDB; 1K2C; 10-JUL-02.
DR PDB; 1LOX; 01-APR-97.
DR PDB; 1TCX; 07-DEC-96.
DR PDB; 1WJE; 16-DEC-98.
DR PDB; 1WJF; 16-DEC-98.
DR PDB; 2BPV; 23-FEB-99.
DR PDB; 2BPW; 23-FEB-99.
DR PDB; 2BPY; 23-FEB-99.
DR PDB; 2BPZ; 23-FEB-99.
DR PDB; 2BPZ; 23-FEB-99.
DR HIV; K02012; POLSBS5.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Approtease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn_1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.

```

```

DR PROSITE; P800141; ASP_PROTEASE; 1.
DR PROSITE; P850175; ASF_PROT_RETROV; 1.
DR AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93 BY SIMILARITY.
SQ SEQUENCE 1015 AA; 115015 MW; F79F0CB4A1A92CBE CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1015;
Best Local Similarity 87.5%; Pred. No. 0.0038;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYVGYMDL 32
DB 339 RKRNPDIVGYMDL 354

RESULT 7
POL_HY1BR STANDARD; PRT; 1015 AA.
AC F03367;
DT 21-JUL-1986 (rel. 01, Created)
DV 01-NOV-1988 (rel. 09, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8509333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Dancos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
RN [2]
RP REVISIONS TO 23-35.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT of two isolates from African patients.";
RL Cell 46:63-74(1986).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=9219341; PubMed=1799632;
RA Spinel1 S., Liu Q.Z., Alzari P.M., Harel P.H., Poljak R.U.;
RT "The three-dimensional structure of the aspartyl protease from the
RT HIV-1 isolate BRU ";
RL Biochimie 73:1391-1396(1991).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02013; -. NOT_ANNOTATED_CDS.
DR PDB; 1HHF; 15-OCT-92.
DR PDB; 1A8G; 15-JUL-98.
DR PDB; 1A94; 16-FEB-99.

```

```
DR PDB: 1D4S; 12-OCT-99.
DR PDB: 1D4Y; 13-OCT-99.
DR PDB: 1DAZ; 03-MAY-00.
DR PDB: 1DIF; 08-MAR-96.
DR PDB: 1HOS; 31-OCT-93.
DR PDB: 1HPO; 21-APR-97.
DR PDB: 1HPS; 15-JAN-95.
DR PDB: 1HPX; 08-MAR-96.
DR PDB: 1HSG; 03-APR-96.
DR PDB: 1HTE; 31-JUL-94.
DR PDB: 1HVL; 30-APR-94.
DR PDB: 1UPJ; 14-OCT-96.
DR PDB: 2UPJ; 14-OCT-96.
DR PDB: 7UPJ; 21-APR-97.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase Zn.
DR InterPro; IPR009007; Pept A acid.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROT_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypeptide, Hydroxylase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93 BY SIMILARITY.
FT STRAND 78 82
FT STRAND 87 92
FT TURN 94 95
FT STRAND 100 102
FT TURN 111 117
FT TURN 118 119
FT STRAND 120 134
FT TURN 135 136
FT STRAND 137 146
FT STRAND 152 153
FT HELIX 155 161
FT TURN 162 162
SQ SEQUENCE 1015 AA; 115031 MW; 164702F074A84394 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1015;
Best Local Similarity 87.5%; Pred. No. 0.0038;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
DB 339 RKQNPDIIVYQYMDL 354
```

```
RA RX MEDLINE=85111157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RT Aids/lymphadenopathy retrovirus.";
RL Nature 313:450-458 (1985).
RN [2]
RP REVISION.
RA Muesing M.A.;
RL Submitted (xxx-1987) to the HIV data bank.
CC -! CATALYTIC ACTIVITY: Specific for a PL residue that is hydrophobic,
CC and PL variable, but often Pro.
CC -! CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -! CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -! PM: cleavage sites that yield the mature proteins remain to be
CC determined.
CC -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K02083; AB59867.1; -.
DR EMBL; X01762; -; NOT ANNOTATED_CDS.
DR PIR; A03967; GNWVVL.
DR PDB; 1A9M; 17-JUN-98.
DR PDB; 1AJV; 20-AUG-97.
DR PDB; 1AJX; 17-SEP-97.
DR PDB; 1G35; 06-JUN-01.
DR PDB; 1GNM; 08-NOV-96.
DR PDB; 1HTG; 31-JUL-94.
DR PDB; 1HVI; 30-APR-94.
DR PDB; 1KIH; 06-MAR-02.
DR PDB; 1NPV; 04-FEB-03.
DR PDB; 1NPW; 04-FEB-03.
DR HIV; K02083; POL$PV22.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase Zn.
DR InterPro; IPR009007; Pept A acid.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polypeptide, Hydroxylase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 69 167
FT ACT SITE 93 93 BY SIMILARITY.
SQ SEQUENCE 1015 AA; 115090 MW; 51529D18E8C2AF89 CRC64;

Query Match 44.6%; Score 75; DB 1; Length 1015;
Best Local Similarity 87.5%; Pred. No. 0.0038;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
DB 339 RKQNPDIIVYQYMDL 354
```

```

RESULT 9
POL_HV1JR STANDARD; PRT: 1007 AA.
ID POL_HV1JR
AC P20875;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL. polyprotein [contains: Protease (Retrovirsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
OS POL.
NCBI_TaxID=11688;
Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
VIRUSES; Retrovirdae; Retroviridae; Lentivirus.
[1]
SEQUENCE FROM N.A.
RA Koyanagi S., Chen I.S.Y.;
RL Submitted (DEC-1988) to the HIV data bank.
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M38429; AAB03745.1; -.
CC HSSP; P03366; 1HMV.
DR HIV; M38429; POL$JRCSE.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; naseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvc; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase.
FT CHAIN 61 159
FT ACT_SITE 85 85 BY SIMILARITY.
SQ SEQUENCE 1007 AA; 114081 MW; B123656C2EP3411A CRC64;
Query Match 44.0%; Score 74; DB 1; Length 1007;
Best Local Similarity 81.2%; Pred. No. 0.0053;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
DB 331 RKNPDIIVYQYMDL 346

```

```

RESULT 10
POL_HV1J2 STANDARD; PRT: 1003 AA.
ID POL_HV1J2
AC P35963;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL. polyprotein [contains: Protease (Retrovirsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
OS POL.
NCBI_TaxID=36377;
Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
VIRUSES; Retrovirdae; Retroviridae; Lentivirus.
[1]
SEQUENCE FROM N.A.
RA MEDLINE=93021387; Pubmed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation."
RL J. Virol. 66:6587-6600(1992).
CC -1- CATALYTIC ACTIVITY: Specific for a PI residue that is hydrophobic,
CC and PI' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M93258; -; NOT_ANNOTATED_CDS.
DR PIR; B44001; B44001.
DR PDB; 1K6C; 06-FEB-02.
DR PDB; 1K6P; 06-FEB-02.
DR PDB; 1K6T; 06-FEB-02.
DR PDB; 1K6V; 06-FEB-02.
DR PDB; 1WJA; 13-MAY-98.
DR PDB; 1WJC; 13-MAY-98.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; naseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvc; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 57 155
FT ACT_SITE 81 81 BY SIMILARITY.
SQ SEQUENCE 1003 AA; 113794 MW; D2F7187FE4EBF49 CRC64;
Query Match 43.5%; Score 73; DB 1; Length 1003;

```

Best Local Similarity 81.2%; Pred. No. 0.0075;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
| : : : : :
Db 327 RKNPDLVIYQYMDL 342

```
RESULT 11
POL_HV1MA STANDARD; PRT; 1002 AA.
AC P04588; Q79582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11697;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RU Cell 46:63-74(1986).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04415; CAA28012.1; -.
CC EMBL; A07116; CAA00619.1; -.
CC PDB; 1HHJ; 31-OCT-93.
CC HIV; K03456; POLSMAL.
CC MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; Rvt; 1.
DR Pfam; PF00078; Rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR AIDS; Polypoteins; Hydrolyase; Aspartyl protease; Endonuclease;
RNase; Transferrase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 56 154
FT ACT_SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113537 MW; 569A29D69AC6AC5 CRC64;
```

Query Match 42.9%; Score 72; DB 1; Length 1002;
Best Local Similarity 81.2%; Pred. No. 0.011;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
| : : : : :
Db 326 RTKNPEIYQYMDL 341

```
RESULT 12
POL_HV1U4 STANDARD; PRT; 1002 AA.
AC P24740;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
(HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91090981; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
Carwell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
diversity from other HIV-1 isolates.";
RU AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- P1M: Cleavage sites that yield the mature proteins remain to be
determined.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62320; AAA75019.1; -.
CC PDB; 1E27; 12-SEP-00.
CC MEROPS; A02.001; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RYase.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; Rvt; 1.
DR Pfam; PF00078; Rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR AIDS; Polypoteins; Hydrolyase; Aspartyl protease; Endonuclease;
RNase; Transferrase; RNA-directed DNA polymerase; 3D-structure.
FT CHAIN 56 154
FT ACT_SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 113537 MW; 569A29D69AC6AC5 CRC64;
```


DR PDB: 3HVT; 15-OCT-94.
DR PDB: 1HVI; 30-APR-94.
DR PDB: 1HVJ; 30-APR-94.
DR PDB: 1HVK; 30-APR-94.
DR PDB: 1HVL; 30-APR-94.
DR PDB: 1HEF; 31-MAY-94.
DR PDB: 1HEG; 31-MAY-94.
DR PDB: 1HWY; 31-MAR-95.
DR PDB: 1HNT; 03-JUN-95.
DR PDB: 1HNV; 10-JUL-95.
DR PDB: 1HPS; 31-AUG-94.
DR PDB: 1HRE; 31-JUL-94.
DR PDB: 1HTE; 31-JUL-94.
DR PDB: 1HNG; 31-JUL-94.
DR PDB: 1SBG; 15-OCT-94.
DR PDB: 1DLO; 01-AUG-96.
DR PDB: 1GNN; 08-NOV-96.
DR PDB: 1GNO; 08-NOV-96.
DR PDB: 1ATV; 20-AUG-97.
DR PDB: 1ATX; 17-SEP-97.
DR PDB: 1MER; 15-APR-98.
DR PDB: 1MES; 15-APR-98.
DR PDB: 1MET; 15-APR-98.
DR PDB: 1MEU; 15-APR-98.
DR PDB: 1BOM; 16-FEB-99.
DR PDB: 1BON; 16-FEB-99.
DR PDB: 1D4H; 26-JUN-02.
DR PDB: 1D4I; 26-JUN-02.
DR PDB: 1DW6; 26-JUL-00.
DR PDB: 1EBK; 26-JUL-00.
DR PDB: 1EBM; 26-JUN-02.
DR PDB: 1EBY; 26-JUN-02.
DR PDB: 1EBZ; 26-JUN-02.
DR PDB: 1EC0; 26-JUN-02.
DR PDB: 1EC1; 26-JUN-02.
DR PDB: 1EC2; 26-JUN-02.
DR PDB: 1EC3; 26-JUN-02.
DR PDB: 1EET; 07-FEB-01.
DR PDB: 1HBV; 10-JUL-95.
DR PDB: 1HH; 10-JUL-95.
DR PDB: 1HPZ; 30-MAY-01.
DR PDB: 1HOU; 30-MAY-01.
DR PDB: 1HXS; 26-MAR-01.
DR PDB: 1IKV; 17-APR-02.
DR PDB: 1IKX; 17-APR-02.
DR PDB: 1J5O; 14-JUN-02.
DR PDB: 1NSY; 28-JAN-03.
DR PDB: 1N6O; 14-JAN-03.
DR PDB: 1NMC; 14-DEC-99.
DR PDB: 1KTD; 12-JAN-00.
DR PDB: 1TVR; 12-MAR-97.
DR PDB: 1UWB; 15-MAY-97.
DR PDB: 3TLH; 03-FEB-00.
DR HIV; M15654; POLSBH102.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR00308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.

Query Match 42.9%; Score 72; DB 1; Length 1015;
Best Local Similarity 92.9%; Pred. No. 0.011;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 19 RNPDIIVYQYMDL 32
Db 341 QNPDIIVYQYMDL 354

RESULT 14
ID POL_SIVCZ STANDARD; PRT; 1027 AA.
AC P17283;
DT 01-AUG-1990 (rel. 15, Created)
DT 01-AUG-1990 (rel. 15, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE POL polypeptide [Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Chimpanzee immunodeficiency virus (SIVcpz) (CIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=90259077; PubMed=2188136;
RA Huet T., Cheyrier R., Meyerhans A., Roelants G., Main-Hobson S.,
RT "Genetic organization of a chimpanzee lentivirus related to HIV-1,"
RL Nature 345:356-359(1990).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -!- MISCELLANEOUS: N-TERMINUS OF POL PROTEIN IS UNCERTAIN POSSIBLY DUE
CC TO A RIBOSOMAL FRAME SHIFT EVENT BETWEEN GAG AND POL.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; X52154; -; NOT ANNOTATED_CDS.
DR PIR; S09984; GNLJSI.
DR HSSP; P03366; 1RVL.
DR HIV; X52154; POLSCPZ.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR00308; Integrase_Zn.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; ivp; 1.
DR Pfam; PF00078; ivt; 1.
DR PROSITE; PS00141; ASP_PROTASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
KW AIDS; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase.
FT CHAIN 1 179
FT ACT_SITE 105 105
FT SEQUENCE 1027 AA; 116930 MW; C6D374E48431BC1 CRC64;

Query Match 42.9%; Score 72; DB 1; Length 1027;
Best Local Similarity 81.2%; Pred. No. 0.011;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 17 RGNPDIIVYQYMDL 32

Db 351 REKNPDITYQYMDL 366

```
RESULT 15
POL_HVIND
ID POL_HVIND STANDARD; PRT; 1002 AA.
AC P18602;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [contains: Protease (retropepsin) (EC 3.4.23.16);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus.";
RL Gene 81:275-284(1989).
CC -1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
CC and P1 variable, but often Pro.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M27323; AAA44869.1; -.
CC DR PIR; JQ0067; GNLJND.
CC DR HSSP; P03366; IHMV.
CC DR HIV; M27323; POL$NDK.
CC DR MEROPS; A02.001; -.
CC DR InterPro; IPR001969; Asprotease_AS.
CC DR InterPro; IPR001037; Integrase_C.
CC DR InterPro; IPR003308; Integrase_Zn.
CC DR InterPro; IPR009007; Pept_A_acid.
CC DR InterPro; IPR001995; Peptidase_A2.
CC DR InterPro; IPR002155; RNaseH.
CC DR InterPro; IPR001584; Rve.
CC DR InterPro; IPR000477; RVTse.
CC DR Pfam; PF00552; Integrase_1.
CC DR Pfam; PF02022; Integrase_Zn; 1.
CC DR Pfam; PF00075; rnaseH; 1.
CC DR Pfam; PF00665; rve; 1.
CC DR Pfam; PF00077; rvp; 1.
CC DR Pfam; PF00078; rvt; 1.
CC DR PROSITE; PS00141; ASP_PROTEASE; 1.
CC DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
CC DR AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
CC Nuclease; Transferase; RNA-directed DNA polymerase.
CC KW CHAIN 56 154
CC FT ACT_SITE 80 80 BY SIMILARITY.
CC SQ SEQUENCE 1002 AA; 113621 MW; 5ED866B8C9896CC1 CRC64;
```

Query Match 42.3%; Score 71; DB 1; Length 1002;
Best Local Similarity 81.2%; Pred. No. 0.015;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 17 RGRNPDIVTYQYMDL 32
Db 326 RKONPEIVTYQYMDL 341

Search completed: July 20, 2004, 06:31:01
Job time : 6.29508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 32.2623 Seconds
(without alignments)
312.953 Million cell updates/sec

Title: US-09-673-166b-275
Perfect score: 168
Sequence: 1 GRQYKANSKPTGTERGRNPDIYQYMDL 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriophage: *
17: sp_archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.0	328	15	Q9EFP6 human immun
2	84	50.0	328	15	Q9EFC0 human immun
3	84	50.0	328	15	Q9EFP4 human immun
4	84	50.0	328	15	Q9EFC3 human immun
5	84	50.0	328	15	Q9EFC8 human immun
6	84	50.0	328	15	Q9EFC9 human immun
7	84	50.0	328	15	Q9EFC8 human immun
8	84	50.0	328	15	Q9EFC8 human immun
9	84	50.0	328	15	Q9EFC3 human immun
10	84	50.0	328	15	Q9EFC7 human immun
11	84	50.0	328	15	Q9EFC4 human immun
12	84	50.0	328	15	Q9EFC4 human immun
13	84	50.0	328	15	Q9EFC1 human immun
14	84	50.0	328	15	Q9EFC6 human immun
15	84	50.0	328	15	Q9EFC2 human immun
16	84	50.0	341	15	Q9Q1L1 human immun

17	82	48.8	220	15	Q37429 human immun
18	82	48.8	220	15	Q37404 human immun
19	82	48.8	308	15	Q805U7 human immun
20	81	48.2	190	15	Q7ZSR6 human immun
21	81	48.2	207	15	Q7SNCO human immun
22	81	48.2	207	15	Q7SNB9 human immun
23	81	48.2	409	15	Q9DPE5 human immun
24	81	48.2	411	15	Q8ALQ7 human immun
25	81	48.2	999	15	Q8UTQ9 human immun
26	80	47.6	85	15	Q9WFI6 human immun
27	80	47.6	328	15	Q9EFC7 human immun
28	80	47.6	330	15	Q8EAX9 human immun
29	79	47.0	163	15	Q9WGP8 human immun
30	79	47.0	190	15	Q7SSJ5 human immun
31	79	47.0	191	15	Q902D9 human immun
32	79	47.0	191	15	Q902D1 human immun
33	79	47.0	192	15	Q9WGP5 human immun
34	79	47.0	193	15	Q8Q229 human immun
35	79	47.0	200	15	Q8AMJ5 human immun
36	79	47.0	201	15	Q9IFC3 human immun
37	79	47.0	203	15	Q93140 human immun
38	79	47.0	206	15	Q7SJZ5 human immun
39	79	47.0	206	15	Q7SJTO human immun
40	79	47.0	206	15	Q7SJ59 human immun
41	79	47.0	207	15	Q7SNH7 human immun
42	79	47.0	207	15	Q7SNG1 human immun
43	79	47.0	207	15	Q7SNC7 human immun
44	79	47.0	210	15	Q8JSP4 human immun
45	79	47.0	210	15	Q8APX9 human immun

ALIGNMENTS

RESULT 1
Q9EFP6 ID Q9EFP6 PRELIMINARY: PRT; 328 AA.
AC Q9EFP6; 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Pol protein (fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C03;
RA MEDLINE=20408577; PubMed=10952598;
RA Bachele L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aufay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.U., Spalding D.R., Hollis G., Abremki K.,
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
Failing Efavirenz Combination Therapy.";
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002129; AAP92284.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspartic-AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.

```
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1
FT CHAIN 100 >99 PROTEASE.
FT CHAIN 328 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37180 MW; 35CEBB2983A8FC07 CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
DB 271 RGRNPDIYQYMDL 286

RESULT 2
Q9EFQ0 PRELIMINARY; PRT; 328 AA.
AC Q9EFQ0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C02;
MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Anticreb. Agents Chemother. 44:2475-2484(2000).
CC -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002125; AAF92280.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; rvp; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37310 MW; 3415A4E02F247FE8 CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
DB 271 RGRNPDIYQYMDL 286
```

```
RESULT 3
Q9EFN4 PRELIMINARY; PRT; 328 AA.
AC Q9EFN4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=091B05;
MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Anticreb. Agents Chemother. 44:2475-2484(2000).
CC -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002141; AAF92296.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; rvp; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PSS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37297 MW; A4704362BCB77CD CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIYQYMDL 32
DB 271 RGRNPDIYQYMDL 286

RESULT 4
Q9EFQ9 PRELIMINARY; PRT; 328 AA.
AC Q9EFQ9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN=041F05;
MEDLINE=20408577; PubMed=10952598;
RA Bachelel L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., AuJay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002116; AAF92271.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00077; rvt; 1.
DR PROSITE: PS00141; ASP_PROT_RETROV; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37181 MW; 776457F54CE35EA4 CRC64;

Query Match
Best Local Similarity 50.0%; Score 84; DB 15; Length 328;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGNPDIYQYMDL 32
Db 271 RGNPDIYQYMDL 286

RESULT 5
Q9EF08 PRELIMINARY; PRT; 328 AA.
AC Q9EF08;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002117; AAF92272.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00077; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37181 MW; 776457F54CE35EA4 CRC64;

```

```

DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR000477; rvtse.
DR Pfam: PF00077; rvt; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37297 MW; A4704362BCEB77CD CRC64;

Query Match
Best Local Similarity 93.8%; Score 84; DB 15; Length 328;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGNPDIYQYMDL 32
Db 271 RGNPDIYQYMDL 286

RESULT 6
Q9EFN9 PRELIMINARY; PRT; 328 AA.
AC Q9EFN9;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL: AY002136; AAF92291.1; -
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR001995; Peptidase A2.
DR InterPro: IPR009007; Pept_A_acid.
DR InterPro: IPR000477; rvtse.
DR Pfam: PF00077; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37297 MW; A4704362BCEB77CD CRC64;

```

```

SQ SEQUENCE 328 AA, 37218 MW, 31F2676CB94F2EB6 CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
   ||:|||||
DB 271 RGRNPDIIVYQYMDL 286

RESULT 7
ID Q9EFN8 PRELIMINARY; PRT; 328 AA.
AC Q9EFN8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C09;
MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy."
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
RL -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002137; AAF92292.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase_A5.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00078; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 1 >99
FT CHAIN 100 >328
FT NON_TER 328 328
FT CHAIN 328 328
FT NON_TER 328 328
SQ SEQUENCE 328 AA, 37252 MW, 36783542BCFF2EAD CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
   ||:|||||
DB 271 RGRNPDIIVYQYMDL 286

RESULT 8
ID Q9EFN3 PRELIMINARY; PRT; 328 AA.
AC Q9EFN3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
```

```

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C06;
MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy."
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
RL -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002142; AAF92297.1; -.
DR HSP; P0369; ICP1.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Asparticase_A5.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR009007; Pept_A_acid.
DR Pfam: PF00078; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolyase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT CHAIN 1 1
FT NON_TER 1 >99
FT CHAIN 100 >328
FT NON_TER 328 328
FT CHAIN 328 328
FT NON_TER 328 328
SQ SEQUENCE 328 AA, 37281 MW, 1B7ED69EB1BD1B0E CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGRNPDIIVYQYMDL 32
   ||:|||||
DB 271 RGRNPDIIVYQYMDL 286

RESULT 9
ID Q9EFN3 PRELIMINARY; PRT; 328 AA.
AC Q9EFN3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=062C07;
MEDLINE=20408577; PubMed=10952598;
RA Bacheler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abramski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
```

```
RT Failing Efavirenz Combination Therapy."
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC EMBL: AY002132; AAF92287.1;
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR009007; Peptidase A2.
DR InterPro: IPR00477; RVTse.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1 1
FT CHAIN 100 >328 PROTEASE.
FT NON_TER 328 REVERSE TRANSCRIPTASE.
SQ SEQUENCE 328 AA; 37282 MW; 36783542BCB7FE8 CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 RGNPDIVITYOMDL 32
Db 271 RGNPDIVITYOMDL 286

RESULT 10
Q9EFN7 PRELIMINARY; PRT; 328 AA.
AC Q9EFN7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=041F01; PubMed=10952598;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheiler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Antay M., Wang X.V., Ellis D., Becker M.F.,
RA Laust A.L., George H.U., Spalding D.R., Hollis G., Adremski K.,
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy."
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC EMBL: AY002138; AAF92293.1;
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR009007; Peptidase A2.
DR InterPro: IPR00477; RVTse.
```

```
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1 1
FT CHAIN 100 >328 PROTEASE.
FT NON_TER 328 REVERSE TRANSCRIPTASE.
SQ SEQUENCE 328 AA; 37282 MW; 36783542BCB7FE8 CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 RGNPDIVITYOMDL 32
Db 271 RGNPDIVITYOMDL 286

RESULT 11
Q9EFP4 PRELIMINARY; PRT; 328 AA.
AC Q9EFP4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=070G07; PubMed=10952598;
RX MEDLINE=20408577; PubMed=10952598;
RA Bacheiler L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Antay M., Wang X.V., Ellis D., Becker M.F.,
RA Laust A.L., George H.U., Spalding D.R., Hollis G., Adremski K.,
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy."
RT Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC EMBL: AY002131; AAF92286.1;
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro: IPR001969; Aspprotease AS.
DR InterPro: IPR009007; Peptidase A2.
DR InterPro: IPR00477; RVTse.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW transferase.
FT NON_TER 1 1
FT CHAIN 100 >328 PROTEASE.
FT NON_TER 328 REVERSE TRANSCRIPTASE.
SQ SEQUENCE 328 AA; 37313 MW; B5241227A9BA238D CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 17 RGNPDIYIYQYMDL 32
DB 271 RGNPDIYIYQYMDL 286

RESULT 12
QYEF04 PRELIMINARY; PRT; 328 AA.
AC Q9EF04;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID=11676;

SEQUENCE FROM N.A.
RX STRAIN=091B03; PubMed=10952598;
RX MEDLINE=20408577; PubMed=10952598;
RX Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA Laout A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002121; AAF92276.1; -
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteinolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Aspprotease_AS.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 1
FT NON_TER 1 >99
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37226 MW; 29AE22FD46F8DAA2 CRC64;

Query Match
Best Local Similarity 93.8%; Score 84; DB 15; Length 328;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGNPDIYIYQYMDL 32
DB 271 RGNPDIYIYQYMDL 286

RESULT 13
QYEF01 PRELIMINARY; PRT; 328 AA.
AC Q9EF01;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
```

```
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=070G04;
RX MEDLINE=20408577; PubMed=10952598;
RX Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA Laout A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002124; AAF92279.1; -
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteinolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Aspprotease_AS.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR009007; Pept_A_acid.
DR Pfam; PF00077; rvt; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW Transferase.
FT CHAIN 1 1
FT NON_TER 1 >99
FT CHAIN 100 >328 REVERSE TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37297 MW; AA704362BCB77CD CRC64;

Query Match
Best Local Similarity 93.8%; Score 84; DB 15; Length 328;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 RGNPDIYIYQYMDL 32
DB 271 RGNPDIYIYQYMDL 286

RESULT 14
QYEF06 PRELIMINARY; PRT; 328 AA.
AC Q9EF06;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID=11676;

SEQUENCE FROM N.A.
RX STRAIN=054G09;
RX MEDLINE=20408577; PubMed=10952598;
RX Bachelet L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Bolling L., Aujay M., Wang X.V., Ellis D., Becker M.F.,
RA Laout A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy.";
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
EMBL: AY002119; AAF92274.1; -
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
```

DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Aspartase A2.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; Tvp; 1.
DR Pfam; PF00078; tvc; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KM Aspartyl protease, Hydrolase; Protease; RNA-directed DNA polymerase;
Transferrase.
KM NON_TER 1 1
FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE_TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37138 MM; E7575063F438782F CRC64;

Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 17 RGRNPDIYQYWMDDL 32
||:|||||
Db 271 RGRNPDIYQYWMDDL 286

RESULT 15
Q9EPR2 PRELIMINARY; PRT; 328 AA.
AC Q9EPR2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pol protein (Fragment).
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=070605;
RX MEDLINE=20408577; PubMed=10952598;
RA Bachevalier L.T., Anton E., Kudish P., Baker D., Bunville J.,
RA Krakowski K., Boiling L., Anjay M., Wang X.V., Ellis D., Becker M.F.,
RA Lasut A.L., George H.J., Spalding D.R., Hollis G., Abremski K.;
RT "Human Immunodeficiency Virus Type 1 Mutations Selected in Patients
RT Failing Efavirenz Combination Therapy".
RL Antimicrob. Agents Chemother. 44:2475-2484(2000).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
CC KNOWN AS THE RETROPEPSIN FAMILY.
DR EMBL; AY002113; AAF92268.1; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR001969; Aspartase A2.
DR InterPro; IPR001995; Peptidase A2.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; Tvp; 1.
DR Pfam; PF00078; tvc; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase;
KW Transferrase.
FT NON_TER 1 1

FT CHAIN 1 >99 PROTEASE.
FT CHAIN 100 >328 REVERSE_TRANSCRIPTASE.
FT NON_TER 328 328
SQ SEQUENCE 328 AA; 37297 MM; A4704362BCB77CD CRC64;
Query Match 50.0%; Score 84; DB 15; Length 328;
Best Local Similarity 93.8%; Pred. No. 0.00082;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 17 RGRNPDIYQYWMDDL 32
||:|||||
Db 271 RGRNPDIYQYWMDDL 286

Search completed: July 20, 2004, 06:33:18
Job time : 34.2623 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 47.7377 Seconds

(without alignments)
189,400 Million cell updates/sec

Title: US-09-673-166B-275

Perfect score: 168
Sequence: 1 GRQYIKANSKFIGITERGRNPDIYQYWDL 32

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq29and4:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	98	58.3	28	2	AA53552
2	92.5	55.1	50	2	AAW06132
3	92.5	55.1	50	2	AAW46447
4	87	51.8	29	2	AA53551
5	86.5	51.5	25	4	AA884738
6	82.5	49.1	50	2	AAW06131
7	81	48.2	29	5	ABB79182
8	78.5	46.7	280	2	AAW71642
9	78.5	46.7	280	2	AAW71363
10	78.5	46.4	137	4	AA898912
11	78	46.4	137	3	AAV82634
12	78	46.4	218	3	AAV49253
13	78	46.4	240	3	AAV49254
14	78	46.4	390	3	AAV49255
15	78	46.4	750	3	AAV92638
16	78	46.4	1003	3	AA869279
17	78	46.4	1445	4	AAE04828
18	77	45.8	213	5	AAU96698
19	77	45.8	999	3	AA869286
20	76.5	45.5	31	2	AAW06129
21	76.5	45.5	31	2	AAV02470
22	76.5	45.5	126	3	AA845490
23	76	45.2	29	2	AA883561
24	76	45.2	37	2	AA85389
25	76	45.2	37	2	AA85383

26	76	45.2	46	5	AAU1430	AAU1430 Synthetic
27	76	45.2	47	2	AA862723	AA862723 LHRH-cont
28	76	45.2	124	3	AA845519	AA845519 Modified
29	76	45.2	124	3	AA845493	AA845493 Modified
30	76	45.2	141	3	AA845499	AA845499 Modified
31	76	45.2	158	3	AA881329	AA881329 TNF-4, a
32	76	45.2	373	4	AA807281	AA807281 Human TNF
33	76	45.2	158	5	AA881048	AA881048 HIV prote
34	76	45.2	750	3	AA92627	AA92627 Mutant hu
35	75	44.6	20	6	AA835186	AA835186 HIV helpe
36	75	44.6	20	6	AA835159	AA835159 HIV CTL e
37	75	44.6	25	3	AA92652	AA92652 PSMep009
38	75	44.6	25	3	AA92651	AA92651 PSMep008
39	75	44.6	32	2	AA926604	AA926604 HIV-dariv
40	75	44.6	32	5	AA883189	AA883189 lipopepti
41	75	44.6	51	6	AA835202	AA835202 HIV epit
42	75	44.6	60	2	AA894762	AA894762 CTL epit
43	75	44.6	60	4	AA898915	AA898915 HIV Pol-2
44	75	44.6	60	5	AA883192	AA883192 HIV cytot
45	75	44.6	109	4	AA820145	AA820145 Growth di

ALIGNMENTS

RESULT 1	AA53552	standard; protein; 28 AA.
ID	AA53552	
XX	AA53552;	
AC	18-JAN-2000	(first entry)
DT		
XX		
DE	Lipopeptide #3.	
XX		
KW	Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;	
KW	electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;	
KW	human immunodeficiency virus; hepatitis B virus; papilloma virus;	
KW	melanoma; malaria; parasite.	
OS	Synthetic.	
OS	Homo sapiens.	
PH		
FT	Key	Location/Qualifiers
FT	Modified-site	1
FT		/note= "contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"
XX		
XX	FR2776926-A1.	
XX		
PD	08-OCT-1999.	
XX		
XX	07-APR-1998;	98FR-00004323.
XX		
PR	07-APR-1998;	98FR-00004323.
XX		
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.	
PA	(CNRS) CNRS CENT NAT RECH SCI.	
PA	(INSP) INST PASTEUR LILLE.	
XX		
PI	Le Gal FA, Guillet JG, Gahery SH, Gras ME, Melnyk O, Tartar A;	
XX		
DR	WPI; 1999-583113/50.	
XX		
PT	New lipopeptide containing lipid regions and two epitopes, all separated	
XX	by peptide spacers that impart hydrophilicity, useful in vaccines.	
XX		
PS	Example 1; Page 9; 35pp; French.	
XX		
CC	The invention relates to the generation of a lipopeptide comprising at	
CC	least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)	
CC	epitope and at least one lipid residue with (i) the epitopes and lipid	
CC	portion and (ii) the epitopes, being separated independently by peptide	

CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAY5301-Y5349 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 28 AA;

Query Match 58.3%; Score 98; DB 2; Length 28;
Best Local Similarity 84.0%; Pred. No. 1e-07;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GROYIKANSKFIGITGRGRNDIV 25
Db 1 GROYIKANSKFIGITGRGRNAGIGI 25

RESULT 2

ID AAM06132 standard; peptide; 50 AA.

AC AAM06132;

DT 07-FEB-1997 (first entry)

DE Anti-cholesterol ester transfer multivalent vaccine peptide.

KW Cholesterol ester transfer protein; CERP; antigen; vaccine;

KW cardiovascular disease; atherosclerosis.

XX Synthetic.

OS Location/Qualifiers

FH Key

FT Region

/label= "T-cell epitope
/note= "T-cell epitope comprises amino acids 830-843 of
tetanus toxoid protein"

FT Region

/label= "B-cell epitope
/note= "B-cell epitope comprises amino acids 350-368 of
rabbit CERP"

FT Region

/label= "B-cell epitope
/note= "B-cell epitope comprises the C-terminal 16 amino
acids of rabbit CERP" involved in neutral lipid binding
or transfer activity"

FT Region

FT Region

XX WO9634888-A1.

PD 07-NOV-1996.

PR 01-MAY-1996; 96WO-US006147.

PR 01-MAY-1995; 95US-00432483.

PA (TCEL-) T CELL SCI INC.

PI Rittershaus CM, Thomas LJ;

XX WPI; 1996-S06103/50.

XX Cholesterol ester transfer protein B cell epitope linked to T cell

XX epitope - used to generate vaccine to regulate CERP activity for

XX decreasing the risk of developing a cardiovascular disease e.g.

XX atherosclerosis.

XX Disclosure; Page 7; 72pp; English.

CC A multivalent vaccine comprises an immunogenic helper T-cell epitope of
CC tetanus toxoid protein covalently linked to the B-cell epitopes of rabbit
CC cholesterol ester transfer protein (CERP) (see also AAM06132). The
CC vaccine elicits an immune response against endogenous CERP activity, and
CC is used to treat or prevent a cardiovascular disease, such as
CC atherosclerosis

SQ Sequence 50 AA;

Query Match 55.1%; Score 92.5; DB 2; Length 50;
Best Local Similarity 60.6%; Pred. No. 1.4e-06;
Matches 20; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 3 QYIKANSKFIGITGRND---IVIYOMDL 32
Db 2 QYIKANSKFIGITGRPPDGRBAVAYREEDI 34

RESULT 3

ID AAM46447 standard; peptide; 50 AA.

AC AAM46447;

DT 17-OCT-2003 (revised)

DT 27-AUG-2003 (revised)

DT 18-MAY-1998 (first entry)

DE CERP B cell epitope/tetanus toxoid construct for a plasmid vaccine.

KW Cholesterol ester transfer protein; CERP; cholesterol ester;

KW high density lipoprotein; HDL; very low density lipoprotein; VLDL;

KW low density lipoprotein; LDL; atherosclerosis; neutral lipid binding;

KW transfer activity; immunogenic; B cell epitope; antibody;

KW DNA plasmid-based vaccine; broad range helper T cell epitope; treatment;

KW cardiovascular disease; chimeric.

XX Clostridium tetani.

OS Clostridium tetani.

OS Chimeric.

OS Chimeric.

PD 06-NOV-1997.

PR 01-MAY-1997; 97WO-US007294.

PR 01-MAY-1996; 96US-00640713.

PR 21-FEB-1997; 97US-00802967.

PA (TCEL-) T CELL SCI INC.

PI Thomas LJ;

XX WPI; 1997-549731/50.

XX N-PSDB; AAV05128.

XX DNA plasmid-based vaccine encodes CERP B cell and helper T cell

XX epitope(s) - used for elevating high density lipoprotein levels, and for

XX treating cardiovascular disease.

XX Claim 8; Page 22; 67pp; English.

XX The present sequence represents a construct containing a helper T cell

XX epitope from the toxoid tetanus protein, and 2 B cell epitopes of rabbit

PT New lipopeptide containing lipid regions and two epitopes, all separated
PT by peptide spacers that impart hydrophilicity, useful in vaccines.

CC The present sequence represents a lipid-tailed protein. The lipid moiety
CC is able to cross the mucosal membrane, and deliver an antigen to the
CC immune system. The lipoprotein is administered to a mucosal membrane to

XX	AA89912 standard; protein; 280 AA.
AC	AA89912;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Synthetic HIV-derived Pol-Nef chimeric protein.
XX	
KW	HIV-specific immunogen; major histocompatibility complex; MHC class I;
KW	MHC class II; antigen presentation; CD8+ response; CD4+ response;
KW	gene therapy; genetic vaccine; anti-HIV; antiviral; antiretroviral; Pol;
KW	Nef; chimeric protein; CTL epitope; cytotoxic T-lymphocyte.
XX	
OS	Synthetic.
OS	Human immunodeficiency virus.
XX	
FH	Key
FT	Location/Qualifiers
FT	Region
FT	/note= "Pol-3 CTL epitope (AA89913) "
FT	Region
FT	/note= "Nef-2 CTL epitope (AA89914) "
FT	Region
FT	/note= "Pol-2 CTL epitope (AA89915) "
FT	Region
FT	/note= "Nef-1 CTL epitope (AA89916) "
FT	Region
FT	/note= "Pol-1 CTL epitope (AA89917) "
XX	
PN	WO200154701-A1.
XX	
PD	02-AUG-2001.
XX	
PF	26-JAN-2001; 2001WO-US002766.
PR	31-JAN-2000; 2000US-0179276P.
XX	
PA	(AVERT) AVENTIS PASTEUR SA.
PA	(AARO-) AARON DIAMOND AIDS RES CENT.
PI	Ho D, Markowitz M, Klein M, El Habib R;
XX	
DR	WPI, 2001-488741/53.
XX	
PT	Permitting cessation of antiviral therapy on HIV-infected patients
PT	undergoing antiviral therapy, useful for treating HIV-infected patients,
PT	by administering nucleic acid based vaccines encoding HIV-specific
PT	immunogens.
XX	
PS	Disclosure; Page 12; 57pp; English.
XX	
CC	The invention relates to a method for permitting the cessation of
CC	antiviral therapy in HIV-infected patients undergoing antiviral therapy
CC	who have a controlled level of viraemia. The method involves
CC	administration of nucleic-acid based vaccines encoding HIV-specific
CC	immunogens to the patient, enabling the expression of the HIV-specific
CC	immunogens in the patient's cells. The HIV-specific antigens are
CC	presented on major histocompatibility complex (MHC) class I and II
CC	molecules, stimulating an HIV-specific CD8+ and CD4+ response. The method
CC	permits the cessation of antiviral therapy on HIV-infected patients
CC	either without virus rebound, with a delayed rebound, or with decreased
CC	post-rebound set point, in addition to achieving immunological control of
CC	persistent infectious virus after discontinuation of antiviral therapy.
CC	The method is useful for people who have lost their CD4+ and/or CD8+ T
CC	cell responses to HIV antigens, such as those who have lost their
CC	proliferative T cell responses to gp120 or p24. The method is
CC	specifically useful for treating persons infected with a lymphotropic or
CC	immune-destroying retroviral infection, or for treating HIV-infected
CC	patients undergoing anti-retroviral therapy and having a viral load of
CC	less than 10000, preferably 1000 viral copies/ml of plasma and a CD4+
CC	cell count of above 300 cells/ml, preferably 500 cells/ml. The present
CC	sequence represents a synthetic protein comprising N- to C-terminally,
CC	the HIV-specific cytotoxic T-lymphocyte (CTL) epitopes Pol-3 (AA89913),
CC	Nef-2 (AA89914), Pol-2 (AA89915), Nef-1 (AA89916) and Pol-1 (AA89917)
CC	

SQ Sequence 280 AA:

Query Match Best Local Similarity 46.7%; Score 78.5; DB 4; Length 280;
Matches 19; Conservative 4; Mismatches 6; Indels 11; Gaps 2

OY 3 QYIKANSKFIGITE-----RGKRPDPVITYQYMDDL 32
| : | | | | : | : |
Db 61 EYFK-NCTLMALFOSWMTKLIEPFKKDNPDIIVTYQMDDL 119

RESULT 11
ID AAY82634
XX AAY82634 standard; peptide; 137 AA.
AC AAY82634;
DX DT 07-AUG-2000 (first entry)
DE Tetanus toxoid T cell epitopes and Der pII B cell epitopes peptide.
KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
KM antiallergic; antiasthmatic; antiinflammatory; dermatological;
KM immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KM atopic dermatitis; acute urticaria; chronic urticaria;
KM gastro-intestinal syndrome; food allergy; oro-pharyngeal syndrome;
KM anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX XX
OS Dermatophagoides pteronyssinus.
OS Clostridium tetani.
O6 Synthetic.
XX WO200006694-A2.
PN 10-FEB-2000.
XX PD 20-JUL-1999; 99WO-BE000092.
PF PR 30-JUL-1998; 98EP-00870167.
PA (UNIC) UCB SA.
PI Saint-Remy J, Jacquemin M;
XX WPI; 2000-422470/36.
XX New compound for prevention and treatment of allergies comprises at least one allergenic determinant recognized by a B cell and at least one antigenic determinant which does not trigger T cell activation.

Claim 8; Page 35; 50pp; English.

The present invention describes a compound (I) for the prevention and/or treatment of allergy. The compound comprises at least one allergen antigenic determinant (i) recognised by a B cell or an antibody secreted by a B cell of a non-atopic individual and at least one antigenic determinant (ii) different from the allergen that triggers T cell activation. (I) has antiallergic, antiasthmatic, antiinflammatory, dermatological and immunosuppressive activities, and can be used in a vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to treat and/or prevent allergies or a disease of allergic origin, especially hypersensitivities. These include rhinitis, sinusitis, bronchial asthma, atopic dermatitis, some forms of acute and chronic urticaria, gastro-intestinal syndromes associated with the ingestion of food allergens, oro-pharyngeal syndrome, anaphylactic reactions associated with drug hypersensitivities and/or a mixture of these. The use of (I) in the treatment of allergic conditions avoids the need for drug treatment, which often causes undesirable side-effects. Also, prior art drug therapies alleviate symptoms, but do not influence their causes, however (I) actually combats the cause of an allergic reaction. The present sequence represents a specifically claimed compound peptide sequence from the present invention


```

XX  W09955730-A2.
XX
XX  04-NOV-1999.
XX
XX  27-APR-1999; 99WO-IB000844.
XX
XX  27-APR-1998; 98GB-00008932.
XX
XX  (CHIR-) CHIRON SPA.
XX
XX  Rappuoli R, Grandi G;
XX
XX  WPI: 2000-023325/02.
XX
XX  N-PSDB; AA231417.
XX
XX  Carrier proteins containing CD4+ epitopes useful for protecting against
XX  diseases caused by encapsulated bacteria.
XX
XX  Disclosure; Fig 8; 76pp; English.
XX
XX  The invention provides carrier proteins comprising at least 5 CD4+ T cell
XX  epitope. The carrier protein comprises at least 1 of N6, N10 or N19. The
XX  carrier protein can be prepared by expressing a vector comprising a
XX  nucleic acid molecule encoding the protein, in a host cell and recovering
XX  the expressed protein. The carrier protein can also be produced by (a)
XX  constructing oligonucleotide molecules that encode peptide epitopes; (b)
XX  annealing the oligonucleotides to form duplexes; (c) introducing the
XX  duplexes into an expression vector; (d) introducing the expression vector
XX  into a host cell; and (e) isolating the fusion protein produced from a
XX  culture of the host cells. The carrier protein can be used as a
XX  protective immunogen in the control of diseases caused by encapsulated
XX  bacteria. The present sequence represents the amino acid sequence of N19
XX  polypeptide carrier protein construct
XX
XX  Sequence 390 AA;
XX
QY  Query Match 46.4%; Score 78; DB 3; Length 390;
    Best Local Similarity 78.9%; Pred. No. 0.0031; 2; Mismatches 0; Gaps 0;
    Matches 15; Conservative 2; Indels 0;
    3 QYIKANSKFIGITERGNP 21
    Db 126 QYIKANSKFIGITEKGSP 144

RESULT 15
AA92638
ID AA92638 standard; protein; 750 AA.
XX
XX  AA92638;
XX
XX  10-AUG-2000 (first entry)
XX
XX  Mutant human prostate specific membrane antigen construct, hPSM3.1.
XX
XX  Prostate specific membrane antigen; immunogenized construct; mutant;
XX  vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX  prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX  Homo sapiens.
XX
XX  Synthetic.
XX
XX  Key Location/Qualifiers
XX  Peptide 21..41
XX  /label= P30
XX  /note= "Foreign epitope"
XX  Peptide 213..227
XX  /label= P2
XX  /note= "Foreign epitope"
XX
XX  WO200020027-A2.

```

```

PD  13-APR-2000.
XX
XX  05-OCT-1999; 99WO-DK000525.
XX
XX  05-OCT-1998; 98DK-00001261.
XX
XX  20-OCT-1998; 98US-0105011P.
XX
XX  (MERI-) M & B BIOTECH AS.
XX
XX  Steinaa L, Nouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX  Gautam A, Birk P, Karlsson G;
XX
XX  WPI: 2000-349917/30.
XX
XX  Inducing immune responses to weakly immunogenic, tumor associated peptide
XX  antigens for the treatment of breast and prostate cancer.
XX
XX  Example 1; Page; 220pp; English.
XX
XX  AA92627-49 are mutant immunogenized human prostate specific membrane
XX  antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX  The immunogenic analogues of PSM can be used in the claimed method as an
XX  autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX  binding regions and cysteine residues involved in disulfide bonds are
XX  preserved in the immunogenized forms. The method is used for inducing
XX  immune responses against weakly immunogenic cell-associated peptide
XX  antigens (PA) such as those associated with cancers (self-proteins), e.g.
XX  human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX  fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX  simultaneous presentation by antigen producing cells (APCs) of the
XX  animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX  group derived from the PA and/or at least 1 B-cell group derived from the
XX  cell-associated PA; and (2) at least 1 first T helper cell group which is
XX  foreign to the animal. Analogues of human PSM, human Her2 and
XX  human/murine FGF8b comprising a substantial part of all known and
XX  predicted CTL and B-cell epitopes of the respective PA and including at
XX  least one foreign T helper epitope are also claimed. The method is used
XX  to treat prostate, prostate/breast or breast cancer when the PA is human
XX  PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX  from the wild type human PSM (AA92619), which appears on pages 184-187
XX  of the specification
XX
XX  Sequence 750 AA;
XX
QY  Query Match 46.4%; Score 78; DB 3; Length 750;
    Best Local Similarity 76.2%; Pred. No. 0.0069; 4; Mismatches 0; Gaps 0;
    Matches 16; Conservative 1; Indels 0;
    1 GRQYIKANSKFIGITERGNP 21
    Db 211 GRQYIKANSKFIGITELYSDF 231

```

Search completed: July 20, 2004, 06:30:28
 Job time : 48.7377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:10:18 ; Search time 11.7049 Seconds

(without alignments)
279.413 Million cell updates/sec

Title: US-09-673-166b-276

Perfect score: 180

Sequence: 1 GRQYIKANSKFRIGTERGRFPVTPQVPLRPMTYK 34

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	47.2	206	1 ASLJVL	nef protein - huma
2	85	47.2	206	1 ASLJVL	nef protein - huma
3	85	47.2	206	2 S03245	nef protein (clone
4	83	46.1	204	2 S24985	nef protein - huma
5	83	46.1	205	1 ASLJH3	nef protein - huma
6	83	46.1	205	2 S43467	nef protein - huma
7	83	46.1	205	2 JC7975	HIV accessory prot
8	83	46.1	206	1 ASLJFV	nef protein (clone
9	83	46.1	206	2 S03244	nef protein - huma
10	83	46.1	206	2 S33986	nef protein (clone
11	83	46.1	206	2 JCS400	nef protein - huma
12	79	43.9	206	2 S25937	nef protein - huma
13	78	43.3	97	2 S54385	nef protein - huma
14	78	43.3	207	1 OQLJNR	nef protein - huma
15	78	43.3	212	1 S03247	nef protein (clone
16	77	42.8	182	2 S03247	nef protein - huma
17	77	42.8	209	2 T01673	nef protein (clone
18	77	42.8	210	1 ASLJ02	nef protein - huma
19	77	42.8	218	1 ASLJBR	nef protein - huma
20	74	41.1	205	1 B44963	nef protein - huma
21	73	40.6	204	2 S03246	nef protein (clone
22	72	40.0	205	1 ASLJIK	nef protein - huma
23	72	40.0	214	1 T44001	nef protein - huma
24	70	38.9	1315	1 BICJTN	nef protein - huma
25	65	36.1	229	1 ASLJMA	nef protein - huma
26	65	36.1	255	1 ASLJGN	nef protein - huma
27	62	34.4	255	1 ASLJ82	nef protein - huma
28	61	33.9	250	2 S54851	nef protein - huma
29	61	33.9	256	1 ASLJH2	nef protein - huma

30	61	33.9	260	1 ASLJ02	nef protein - huma
31	60	33.3	240	2 S24572	nef protein - huma
32	59	32.8	257	2 S12160	nef protein - huma
33	58	32.2	226	2 S46353	nef protein - huma
34	56	31.1	251	2 S54867	nef protein - huma
35	56	31.1	255	2 S53099	nef protein - huma
36	55	30.6	238	2 S61205	nef protein - huma
37	55	30.6	239	2 S54852	nef protein - huma
38	55	30.6	244	2 S61208	nef protein - huma
39	55	30.6	246	2 S54853	nef protein - huma
40	55	30.6	250	2 S54849	nef protein - huma
41	55	30.6	250	2 S61207	nef protein - huma
42	54.5	30.3	3430	1 GNMWV	genome polypeptide
43	52	28.9	209	2 S61204	nef protein - huma
44	52	28.9	211	1 ASLJMS	nef protein - huma
45	52	28.9	239	2 S61206	nef protein - huma

ALIGNMENTS

RESULT 1

ASLJVL
nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, 12)

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999

C:Accession: A04006

R:ARYA, S.K.; Gallo, R.C.

Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986

A:Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity of

A:Reference number: A94093; MUID:8617573; PMID:3008154

C:Accession: A04006

A:Molecule type: DNA

A:Residues: 1-206 <ARY>

A:Cross-References: EMBL:M11840; NID:G328453; PIDN:AAA45001.1; PID:G328458

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match

Best Local Similarity 47.2%; Score 85; DB 1; Length 206;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRPPVTPQVPLRPMTYK 34

DB 64 ERVGFVTPQVPLRPMTYK 82

RESULT 2

ASLJVL
nef protein - human immunodeficiency virus type 1 (isolate LV)

N:Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

R:Musning, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Laszky, L.A.; Capon, D.J.

Nucleic Acids Res. 13, 450-458, 1985

A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi

A:Reference number: A93355; MUID:8511157; PMID:2982104

C:Accession: A04007

A:Molecule type: DNA

A:Residues: 1-206 <MUB>

A:Cross-References: GB:X02083; NID:G555008; PIDN:AA559874.1; PID:G328560

C:Genetics:

A:Gene: nef; 3'-orf; orf-F

C:Superfamily: AIDS nef protein

C:Keywords: AIDS; immunodeficiency

Query Match

Best Local Similarity 47.2%; Score 85; DB 1; Length 206;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMTYK 34

Db 64 EKVGFPVTPQVPLRPMTYK 82

RESULT 3
nef protein (clone HXB3) - human immunodeficiency virus type 1

N/Alternate names: 3'-orf protein

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Sep-1999

C/Accession: S03245

R/Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, R.P.; Livak, K.J.; Petteway
Nucleic Acids Res. 13, 8219-8229, 1985

A/Title: Polymorphism of the 3' open reading frame of the virus associated with the acqy
A/Reference number: S03244; PMID:86067228; PMID:2999715

A/Accession: S03245

A/Molecule type: DNA

A/Residues: 1-206 <RAT>
A/Cross-references: EMBL:X03198; NID:g61556; PIDN:CAA26947.1; PID:g61557

C/Genetics:

A/Gene: nef; 3'-orf; orf-F
C/Superfamily: AIDS nef protein

Query Match 47.2%; Score 85; DB 2; Length 206;

Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMTYK 34

Db 64 EKVGFPVTPQVPLRPMTYK 82

RESULT 4

nef protein - human immunodeficiency virus type 1

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999

C/Accession: S24985

R/Harris, M.; Hsiolop, S.; Patsilinos, P.; Neil, J.C.
submitted to the EMBL Data Library, November 1991

A/Description: In vivo derived HIV-1 NEF genes are heterogeneous and lack detectable nuc
A/Reference number: S24985

A/Accession: S24985

A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-204 <HAR>
A/Cross-references: EMBL:X63044; NID:g60123; PIDN:CAA44770.1; PID:g60124

C/Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 204;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 66 FPVTPQVPLRPMTYK 80

RESULT 5

nef protein - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

N/Alternate names: 3'-orf protein; orf-F protein

C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)

C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 31-Jan-1997

C/Accession: A04005
R/Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dor
nberger, J.A.; Papas, T.S.; Garbayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

Nature 313, 277-284, 1985
A/Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A/Reference number: A93353; PMID:85111123; PMID:2578615

A/Accession: A04005
A/Molecule type: DNA

A/Residues: 1-205 <RAT>
C/Genetics:

A/Gene: nef; 3'-orf; orf-F
C/Superfamily: AIDS nef protein
C/Keywords: AIDS; immunodeficiency

Query Match 46.1%; Score 83; DB 1; Length 205;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 68 FPVTPQVPLRPMTYK 82

RESULT 6

nef protein - human immunodeficiency virus type 1

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 03-May-1996

C/Accession: S43467

R/Freund, J.; Kellner, R.; Houthaeve, T.; Kalbitzer, H.R.
Eur. J. Biochem. 221, 811-819, 1994

A/Title: Stability and proteolytic domains of Nef protein from human immunodeficiency vi
A/Reference number: S43467; PMID:94229079; PMID:8174561

A/Accession: S43467
A/Status: preliminary

A/Molecule type: protein
A/Residues: 1-205 <PRE>
C/Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 67 FPVTPQVPLRPMTYK 81

RESULT 7

HIV accessory protein Nef - mouse

C/Species: Mus musculus (house mouse)
C/Date: 25-Aug-2003 #sequence_revision 25-Aug-2003 #text_change 15-Sep-2003

C/Accession: JC7975

R/Seco, P.; Corella, D.; Santoro, C.
Biochem. Biophys. Res. Commun. 305, 1061-1066, 2003

A/Title: Selection of peptides with affinity for the N-terminal domain of GATRA-1: Identifi
A/Reference number: JC7975; PMID:12767938

A/Accession: JC7975
A/Molecule type: DNA

A/Residues: 1-205 <SEC>
C/Comment: This protein, an accessory viral protein, acts as a nuclear regulatory factor
binds p53 via its amino-terminal domain and protects cell against p53-mediated apoptosis

C/Keywords: GATRA-1; Nef

Query Match 46.1%; Score 83; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMTYK 34

Db 68 FPVTPQVPLRPMTYK 82

RESULT 8

nef protein - human immunodeficiency virus type 1 (isolate LAV-1a)

N/Alternate names: 3'-orf protein; orf-F protein

C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A04008; S14609
R:Main-Hosson, S.; Sonigo, P.; Damos, O.; Cole, S.; Alizon, M.
Cell 40, 9-17, 1985
A:Title: Nucleotide sequence of the AIDS virus, LAV.
A:Reference number: A90866; MUID:8509333; PMID:2981635
A:Accession: A04008
A:Molecule type: DNA
A:Residues: 1-206 <MAI>
A:Cross-references: GB:K02013; NID:G326417; PIDN:AA59752.1; PID:G326425
A:Experimental source: isolate LAV-1a
R:Ciccarilli, R. B.
submitted to the EMBL Data Library, March 1991
A:Reference number: S14607
A:Accession: S14609
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <CIC>
A:Cross-references: EMBL:X58780; NID:G60113; PIDN:CAA1585.1; PID:G60114
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein
C:Keywords: AIDS; immunodeficiency; phosphoprotein

Query Match 46.1%; Score 83; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMYTK 34
DB 68 PPVTPQVPLRPMYTK 82

RESULT 9
S03244
nef protein (clone HXB2) - human immunodeficiency virus type 1
N:Alternate names: 3'-orf protein
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 30-Sep-1993
C:Accession: S03244
R:Ratner, L.; Starcich, B.; Josephs, S.F.; Hahn, B.H.; Reddy, E.P.; Livak, K.J.; Petteway
Nucleic Acids Res. 13, 8219-8229, 1985
A:Title: Polymorphism of the 3' open reading frame of the virus associated with the acqu
A:Reference number: S03244; MUID:86067228; PMID:2999715
A:Accession: S03244
A:Molecule type: DNA
A:Residues: 1-206 <RAT>
A:Cross-references: EMBL:X03187
A:Note: the in-frame stop codon at residue 124, shown as 'X', may be suppressed
C:Genetics:
A:Gene: nef; 3'-orf; orf-F
C:Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMYTK 34
DB 68 PPVTPQVPLRPMYTK 82

RESULT 10
S33986
nef protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
C:Accession: S33986
R:Carlini, F.
submitted to the EMBL Data Library, November 1991
A:Reference number: S33979

A:Accession: S33986
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-206 <CAR>
A:Cross-references: EMBL:Z11530; NID:G60192; PIDN:CAA7629.1; PID:G60200
C:Superfamily: AIDS nef protein

Query Match 46.1%; Score 83; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMYTK 34
DB 68 PPVTPQVPLRPMYTK 82

RESULT 11
JC5400
nef protein - Human immunodeficiency virus type 1, HIV-1
C:Species: Human immunodeficiency virus type 1, HIV-1
C:Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 31-Oct-1997
C:Accession: JC5400
R:Macreadie, I.G.; Lowe, M.G.; Curtain, C.C.; Hewish, D.; Azad, A.A.
Biochem. Biophys. Res. Commun. 232, 707-711, 1997
A:Title: Cytotoxicity resulting from addition of HIV-1 Nef N-terminal peptides to yeast
A:Reference number: JC5400; MUID:97271389; PMID:9126340
A:Accession: JC5400
A:Molecule type: protein
A:Residues: 1-206 <MAC>
C:Comment: The amino-terminal part possesses membrane-perturbing and fusogenic activities.
C:Superfamily: AIDS nef protein
F:2-206/Product: nef protein #status predicted <MAT>

Query Match 46.1%; Score 83; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMYTK 34
DB 68 PPVTPQVPLRPMYTK 82

RESULT 12
S25937
nef protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 25-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 20-Sep-1999
C:Accession: S25937
R:Guo, H.G.; Chernann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher, H.
Nature 349, 745-746, 1991
A:Title: Sequence analysis of original HIV-1.
A:Reference number: S25937; MUID:91156044; PMID:2000145
A:Accession: S25937
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: nucleic acid
A:Residues: 1-206 <GUO>
A:Cross-references: EMBL:X57465; NID:G60217; PIDN:CAA40702.1; PID:G60218
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C:Superfamily: AIDS nef protein

Query Match 43.9%; Score 79; DB 2; Length 206;
Best Local Similarity 78.9%; Pred. No. 0.00094;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
DB 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 13
S54385
nef protein - human immunodeficiency virus type 1 (fragment)

C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Sep-1999
C;Accession: S54385
R;Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
A;Reference number: S54377
A;Accession: S54385
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-97 <THE>
A;Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45371.1; PID:g555045
C;Superfamily: AIDS nef protein

Query Match 43.3%; Score 78; DB 2; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.00056;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVPLRPMTYK 34
:| |||||
Db 64 SEEVGFVPRPQVPLRPMTYK 83

RESULT 14

OOLJND

nef protein - human immunodeficiency virus type 1 (isolate NDK)

N;Alternate names: 3'-orf protein; orf-F protein

C;Species: human immunodeficiency virus type 1, HIV-1

A;Note: host Homo sapiens (man)

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C;Accession: J00068

R;Spite, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;

Gene 81, 275-284, 1989

A;Title: Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the human immuno

A;Reference number: J00065; MUID:90034200; PMID:2806917

A;Accession: J00068

A;Molecule type: DNA

A;Residues: 1-207 <SPI>

A;Cross-references: GB:M27323; NID:g328154; PIDN:AAA44874.1; PID:g328163

C;Genetics:

A;Gene: nef

C;Superfamily: AIDS nef protein

C;Keywords: AIDS; immunodeficiency

Query Match 43.3%; Score 78; DB 1; Length 207;

Best Local Similarity 75.0%; Pred. No. 0.0013;

Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVPLRPMTYK 34
:| |||||
Db 64 SEEVGFVPRPQVPLRPMTYK 83

RESULT 15

OOLJZR

nef protein - human immunodeficiency virus Zr-6

N;Alternate names: 3'-orf protein; orf-F protein

C;Species: human immunodeficiency virus Zr-6

C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C;Accession: F26192

R;Srinivasan, A.; Anand, R.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu

Gene 52, 71-82, 1987

A;Title: Molecular characterization of human immunodeficiency virus from Zaïre: nucleoti

A;Reference number: A26192; MUID:87248097; PMID:3036660

A;Accession: F26192

A;Molecule type: DNA

A;Residues: 1-212 <SRI>

A;Cross-references: GB:K03458; GB:M16322; NID:g329398; PIDN:AAA45381.1; PID:g329404

C;Genetics:

A;Gene: nef; 3'-orf; orf-F

C;Superfamily: AIDS nef protein

C;Keywords: AIDS; immunodeficiency

Query Match 43.3%; Score 78; DB 1; Length 212;

Best Local Similarity 75.0%; Pred. No. 0.0013;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVPLRPMTYK 34
:| |||||
Db 69 SEEVGFVPRPQVPLRPMTYK 88

Search completed: July 20, 2004, 06:34:10
Job time : 11.7049 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:17 ; Search time 6.68852 Seconds

(without alignments)
264.690 Million cell updates/sec

Title: US-09-673-166B-276

Perfect score: 180

Sequence: 1 GROYIKANSKFIGITERGRFPVTPQVPLRPMTYK 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	47.2	206	NEF_HV112	P04324 human immun
2	85	47.2	206	NEF_HV112	P04324 human immun
3	83	46.1	123	NEF_HV1B1	P03405 human immun
4	83	46.1	123	NEF_HV1B1	P03405 human immun
5	83	46.1	123	NEF_HV1B1	P03405 human immun
6	83	46.1	123	NEF_HV1B1	P03405 human immun
7	83	46.1	123	NEF_HV1B1	P03405 human immun
8	83	46.1	123	NEF_HV1B1	P03405 human immun
9	83	46.1	123	NEF_HV1B1	P03405 human immun
10	83	46.1	123	NEF_HV1B1	P03405 human immun
11	83	46.1	123	NEF_HV1B1	P03405 human immun
12	83	46.1	123	NEF_HV1B1	P03405 human immun
13	83	46.1	123	NEF_HV1B1	P03405 human immun
14	83	46.1	123	NEF_HV1B1	P03405 human immun
15	83	46.1	123	NEF_HV1B1	P03405 human immun
16	83	46.1	123	NEF_HV1B1	P03405 human immun
17	83	46.1	123	NEF_HV1B1	P03405 human immun
18	83	46.1	123	NEF_HV1B1	P03405 human immun
19	83	46.1	123	NEF_HV1B1	P03405 human immun
20	83	46.1	123	NEF_HV1B1	P03405 human immun
21	83	46.1	123	NEF_HV1B1	P03405 human immun
22	83	46.1	123	NEF_HV1B1	P03405 human immun
23	83	46.1	123	NEF_HV1B1	P03405 human immun
24	83	46.1	123	NEF_HV1B1	P03405 human immun
25	83	46.1	123	NEF_HV1B1	P03405 human immun
26	83	46.1	123	NEF_HV1B1	P03405 human immun
27	83	46.1	123	NEF_HV1B1	P03405 human immun
28	83	46.1	123	NEF_HV1B1	P03405 human immun
29	83	46.1	123	NEF_HV1B1	P03405 human immun
30	83	46.1	123	NEF_HV1B1	P03405 human immun
31	83	46.1	123	NEF_HV1B1	P03405 human immun
32	83	46.1	123	NEF_HV1B1	P03405 human immun
33	83	46.1	123	NEF_HV1B1	P03405 human immun

34	61	33.9	230	NEF_SIVAI	P27970 simian immu
35	61	33.9	256	NEF_HV2RO	P04600 human immu
36	61	33.9	260	NEF_HV2CA	P24103 human immu
37	60	33.3	240	NEF_HV2D2	P18829 human immu
38	59	32.8	257	NEF_HV2BE	P18092 human immu
39	59	32.8	257	NEF_HV2D1	P17753 human immu
40	56	31.1	261	NEF_SIVSP	P19501 simian immu
41	54.5	30.3	3430	POIG_MNV	P06935 w genome po
42	53.5	29.7	1544	ARHC_HUMAN	Q98255 homo sapien
43	52	28.9	207	NEF_SIVMK	P05861 simian immu
44	52	28.9	263	NEF_SIVML	P12482 simian immu
45	52	28.9	309	NEF_SIVS4	

ALIGNMENTS

RESULT 1
ID NEF_HV112 STANDARD; PRT; 206 AA.
AC P04324;

DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (clone 12) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11679;

RP SEQUENCE FROM N.A.
RX MEDLINE=8617573; PubMed=308154;
RA Arya S.K., Gallo R.C.;
RT "Three novel genes of human T-lymphotropic virus type III: immune
RT reactivity of their products with sera from acquired immune
RT deficiency syndrome patients."
RL Proc. Natl. Acad. Sci. U.S.A. 83:2209-2213 (1986).
RN [2]

RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.

RX MEDLINE=86039140; PubMed=3118220;
RA Guy B., Kiely M., Riviere Y., Le Peuch C., Dotti K., Girard M.,
RT Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RT an oncogene product."
RL Nature 330:266-269 (1987).

CC -! FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M1840; AAA5001.1; -
DR PIR: A04006; ASL112.
DR PDB: 1ZEC; 07-JAN-98.
DR HIV; M1840; NEFSPV12.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; Myristate; GTP-Binding; 3D-structure; Lipoprotein.
FT LIPID
SQ SEQUENCE 206 AA; 23366 MW; 218FSB2980F9A46 CRC64;

Query Match 47.2%; Score 85; DB 1; Length 206;
Best Local Similarity 84.2%; Pred. No. 3.1e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

0Y 16 ERGRFPVTPQVPLRPMTYK 34
|: |||||||||||||||||

```

Db      64  EKVGFPTPOVPLRPMYK 82

RESULT 2
NEF_HV1PV STANDARD; PRT; 206 AA.
AC P03405;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11700;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111157; PubMed=2982104;
RA Mueing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RA "Nucleic acid structure and expression of the human
RA AIDS/lymphadenopathy retrovirus.";
RT Nature 313:450-458(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., Le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RA "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RA an oncogene product.";
RT Nature 330:266-269(1987).
RN [2]
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02083; AAB59874.1; -
DR EMBL; X01762; -; NOT_ANNOTATED_CDS.
DR PIR; A04007; ASLJVL.
DR HSSP; P03406; 1EFN.
DR HIV; K02083; NEFSPV22.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
KW AIDS; Myristate; 2 N-myristoyl glycine (in host).
FT LIPID 2
SQ SEQUENCE 206 AA; 23352 MW; EDE64281A17C6735 CRC64;

Query Match 47.2%; Score 85; DB 1; Length 206;
Best Local Similarity 84.2%; Pred. No. 3.1e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      16  ERGRFPVTPVPLRPMYK 34
      1: |||||
      64  EKVGFPTPOVPLRPMYK 82

RESULT 3
NEF_HV1B1 STANDARD; PRT; 123 AA.
AC P03404;
DT 21-JUN-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (BHI0 isolate) (HIV-1).

```

```

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Lyvak K.J., Starcich B.R.,
RA Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Retaway S.R. Jr., Pearson M.L.,
RA Lauenerberger J.A., Papas T.S., Chayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RN Nature 313:277-284(1985).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., Le Peuch C., Dott K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RA "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RA an oncogene product.";
RT Nature 330:266-269(1987).
RN [2]
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -1- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A
CC STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES
CC (210 AA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15654; AAA44206.1; -
DR PIR; A04005; ASLJH3.
DR HIV; M15654; NEFSBH102.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; Myristate; GTP-binding; Lipoprotein.
KW AIDS; Myristate; 2 N-myristoyl glycine (in host).
FT LIPID 2
SQ SEQUENCE 123 AA; 13606 MW; 0811735345F0EB8B CRC64;

Query Match 46.1%; Score 83; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      20  FVPTPOVPLRPMYK 34
      1: |||||
      68  FVPTPOVPLRPMYK 82

RESULT 4
NEF_HV1H2 STANDARD; PRT; 123 AA.
AC P04601; 009780;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8729196; PubMed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RN AIDS Res. Hum. Retroviruses 3:57-69(1987).

```

```

RN      12]
RP      REVISIONS.
RA      Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
OC      Gallo R.C., Wong-Staal F.;
RL      Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN      13]
RP      POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX      MEDLINE=86039140; PubMed=3118220;
RA      Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dotti K., Girard M.,
RT      Montagnier L., Lecocq J.-P.,
RT      "HIV F/3'orf encodes a phosphorylated GTP-binding protein resembling
RL      an oncogene product.";
RL      Nature 330:266-269(1987).
CC      -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC      activities. It seems to down-regulate the CD4(T4) antigen.
CC      -!- MISCELLANEOUS: IN THIS ISOLATE A MUTATION IN POSITION 124 ADDS A
CC      STOP CODON AND THE PROTEIN IS MUCH SHORTER THAN IN OTHER ISOLATES
CC      (210 AA).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sb.ch/announce/
CC      or send an email to license@isb-sb.ch).
CC      -----
DR      EMBL; K03455; AAB50263.1; -.
DR      EMBL; AF033819; AAC82597.1; -.
DR      HIV; K03455; NEFSXB2.
DR      InterPro; IPR001358; HIV_Nef.
DR      Pfam; PF00469; F-protein; 1.
DR      ProDom; PD000031; HIV_Nef; 1.
KW      AIDS; Myristate; GTP-binding; Phosphorylation; Lipoprotein.
FT      LIPID 2
FT      MOD RES 15 2
FT      MOD RES 13692 MW; B5007753CCDD244CF CRC64;
SQ      SEQUENCE 123 AA; 13692 MW; B5007753CCDD244CF CRC64;

Query Match 46.1%; Score 83; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 PPVTPQVPLRPWTYK 34
Db      68 PPVTPQVPLRPMTYK 82
      |||||
      |||||

RESULT 5
ID      NEF_HY1B8 STANDARD; PRT; 205 AA.
AC      P05855;
DT      01-NOV-1988 (Rel. 09, Created)
DT      01-NOV-1988 (Rel. 09, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN      NEF.
OS      Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX      NCBI_TaxId=11684;
[1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=8511123; PubMed=257615;
RA      Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA      Josephs S.F., Doran E.R., Ratajski J.A., Whitehorn E.A.,
RA      Baumeister K., Ivanoff L., Pettway S.R. Jr., Pearson M.L.,
RA      Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA      Wong-Staal F.;
RT      "Complete nucleotide sequence of the AIDS virus, HIV-III.";
RL      Nature 313:277-284(1985).
[2]
RP      POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX      MEDLINE=86039140; PubMed=3118220;
RA      Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dotti K., Girard M.

```

```

RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
RL an oncogene product.";
CC Nature 330:266-269(1987).
CC -I- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL_KO2011; -, NOT_ANNOTATED_CDS.
DR HSSP_P03406; IEFN.
DR HIV_KO2011; NEFSB8.
DR InterPro: IPRO01558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
DR AIDS_Myristate; GTP-Binding, Lipoprotein.
FT LPID 2
SQ SEQUENCE 205 AA; 23305 MW; 8EC12F650DD11 CRC64;
    N-myristoyl glycine (in host).

Query Match          46.1%; Score 83; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 6,2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 PPVTPQVPLRPMYTK 34
Db       68 PPVTPQVPLRPMYTK 82
        |||||
AC P03406; STANDARD; PRT; 206 AA.
ID_NEF_HV1BR
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1), and
OS Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11666, 11698;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate BRU.
RX MEDLINE=85099333; PubMed=2981635;
RX Main-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PNL4-3).
RC STRAIN=Isolate New York-5;
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
RL Submitted (JUN-1988) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RX Guy B., Kieny M.-P., Riviere Y., Le Peuch C., Dotti K., Girard M.,
RA Montagnier L., Lecocq J.-P.;
RT "HIV F/3' orf encodes a phosphorylated GNP-binding protein resembling
RT an oncogene product.";
RL Nature 330:266-269(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 54-205 IN COMPLEX WITH NEF.
RX MEDLINE=96279837; PubMed=8681387;
RX Lee C.H., Sakaela K., Mirza U.A., Chait B.T., Kurtyan J.;
RT "Crystal structure of the conserved core of HIV-1 Nef complexed with
RT a Src family SH3 domain.";
```

```

RL Cell 85:931-942(1996).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02013; AAB59752.1; -.
CC EMBL; M19921; AAA44993.1; -.
CC EMBL; A04321; CAA00353.1; -.
CC PIR; A04008; ASLJFV.
CC PDB; 1EFN; 11-JAN-97.
CC PDB; 1AVZ; 25-MAR-98.
CC PDB; 1AVZ; 25-MAR-98.
CC PDB; 1OA4; 26-MAY-99.
CC DR HIV; K02013; NEFSHRU.
CC DR HIV; M19921; NEFSML43.
CC DR InterPro; IPR001558; HIV_Nef.
CC DR Pfam; PF00469; F-protein; 1.
CC DR ProDom; PD000031; HIV_Nef; 1.
CC DR AIDS; Myristate; GTP-binding; Phosphorylation; 3D-structure;
CC Lipoprotein.
CC KW Lipid
CC FT MOD RES 2 2 N-myristoyl glycine (in host).
CC FT VARIANT 15 15 PHOSPHORYLATION (BY PKC).
CC FT VARIANT 11 11 V -> I (IN CLONE PML4-3).
CC FT VARIANT 15 15 T -> A (IN CLONE PML4-3).
CC FT VARIANT 33 33 A -> V (IN CLONE PML4-3).
CC FT VARIANT 51 51 T -> N (IN CLONE PML4-3).
CC FT HELIX 81 93
CC FT TURN 94 94
CC FT STRAND 98 98
CC FT TURN 99 99
CC FT STRAND 101 101
CC FT HELIX 104 118
CC FT STRAND 127 127
CC FT STRAND 134 134
CC FT STRAND 136 136
CC FT TURN 139 140
CC FT STRAND 143 147
CC FT STRAND 181 185
CC FT HELIX 187 190
CC FT TURN 191 191
CC FT HELIX 194 198
CC FT HELIX 200 202
CC SQ SEQUENCE 206 AA; 23342 MW; 77453FC80B6004F2 CRC64;

Query Match 46.1%; Score 83; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPYTPQVPLRPMYK 34
DB 68 FPYTPQVPLRPMYK 82

RESULT 7
NEF_HV11W STANDARD; PRT; 206 AA.
AC Q70627;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DR Negative Factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirinae.
OX NCBI_TaxID=82834;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
RL infected with HIV type 1 (HIV type IIB)".
RN AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RP STRUCTURE BY NMR OF 56-206.
RX MEDLINE=97337445; PubMed=9194185;
RA Grzesiek S., Bax A., Hu U.S., Kaufman J., Palmer I., Stahl S.J.,
RA Tjandra N., Wingfield P.T.;
RT "Refined solution structure and backbone dynamics of HIV-1 Nef.";
RL Protein Sci. 6:1248-1263(1997).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U12055; AAA76691.1; -.
CC PDB; 2NEF; 07-JUL-97.
CC DR InterPro; IPR001558; HIV_Nef.
CC DR Pfam; PF00469; F-protein; 1.
CC DR ProDom; PD000031; HIV_Nef; 1.
CC KW AIDS; Myristate; GTP-binding; 3D-structure; Lipoprotein.
CC FT Lipid
CC FT SEQUENCE 206 AA; 23414 MW; ED81F68F6B61278E CRC64;

Query Match 46.1%; Score 83; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPYTPQVPLRPMYK 34
DB 68 FPYTPQVPLRPMYK 82

RESULT 8
NEF_HV122 STANDARD; PRT; 97 AA.
AC P12478;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DR Negative Factor (F-protein) (27 kDa protein) (3'ORF) (Fragment).
GN NEF.
OS Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirinae.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22639; AAA45371.1; -.
CC DR PIR; S54385; S54385.
CC DR HIV; M22639; NEFS2226.

```


DR InterPro: IPR001558; HIV_Nef.
 DR Pfam: PF00469; F-protein-1.
 DR Prodom: PD000031; HIV_Nef; 1.
 KM AIDS; Myristate; GTP-binding; Lipoprotein.
 FT Lipid 2 N-myristoyl glycine (in host) (By
 FT NON TER 97 97 similarity).
 SQ SEQUENCE 97 AA; 10592 MW; F8529A06B670AE9 CRC64;

Query Match 43.3%; Score 78; DB 1; Length 97;
 Best Local Similarity 75.0%; Pred. No. 0.00015;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVLRPMTYK 34
 : |||||
 Db 64 SEGVGFVRPQVLRPMTYK 83

RESULT 9
 ID NEF_HV1ND STANDARD; PRT; 207 AA.
 AC P18E01;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
 GN NEF.
 OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 OK NCBI_TaxID=11695;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90034200; PubMed=2806917;
 RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
 RA Hampe A., Chermann J.C.;
 RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
 RT human immunodeficiency virus.",
 RL Gene 81:275-284(1989).
 CC CC
 CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
 CC activities. It seems to down-regulate the CD4(T4) antigen.
 CC -!- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
 CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
 CC CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. By non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC CC
 CC EMBL; M27323; AAA44874.1; -;
 DR PIR; J00068; Q0LND.
 DR HSSP; P03406; IEFN.
 DR HIV; M27323; NEFSNDK.
 DR InterPro: IPR001558; HIV_Nef.
 DR Pfam: PF00469; F-protein; 1.
 DR Prodom: PD000031; HIV_Nef; 1.
 KM AIDS; Myristate; GTP-binding; Lipoprotein.
 FT Lipid 2 N-myristoyl glycine (in host) (By
 FT NON TER 97 97 similarity).
 SQ SEQUENCE 207 AA; 23748 MW; 09036C2F81D5D5E CRC64;

Query Match 43.3%; Score 78; DB 1; Length 207;
 Best Local Similarity 75.0%; Pred. No. 0.00034;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVLRPMTYK 34
 : |||||
 Db 64 SEGVGFVRPQVLRPMTYK 83

RESULT 10

NEF_HV1Z6
 ID NEF_HV1Z6 STANDARD; PRT; 212 AA.
 AC P04602;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
 GN NEF.
 OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 OK NCBI_TaxID=11708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87248097; PubMed=3036660;
 RA Sriivasan A., Anand R., York D., Ranganathan P., Feorino P.,
 RA Schocherman G., Curran J., Kalyanaraman V.S., Luciw P.A.,
 RA Sanchez-Pescador R.;
 RT "Molecular characterization of human immunodeficiency virus from
 RT Zaire: nucleotide sequence analysis identifies conserved and variable
 RT domains in the envelope gene";
 RL Gene 52:71-82(1987).
 CC CC
 CC [2]
 CC POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
 CC MEDLINE=88039140; PubMed=3118220;
 CC Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dott K., Girard M.,
 CC Montagnier L., Lecocq J.-P.;
 CC "HIV F/3' orf encodes a phosphorylated GTP-binding protein resembling
 CC an oncogene product";
 CC Nature 330:266-269(1987).
 CC CC
 CC -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
 CC activities. It seems to down-regulate the CD4(T4) antigen.
 CC CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. By non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC CC
 CC EMBL; K03458; AAA45381.1; -;
 DR PIR; F26192; Q0LZR.
 DR HSSP; P03406; IEFN.
 DR HIV; K03458; NEFS26.
 DR InterPro: IPR001558; HIV_Nef.
 DR Pfam: PF00469; F-protein; 1.
 DR Prodom: PD000031; HIV_Nef; 1.
 KM AIDS; Myristate; GTP-binding; Lipoprotein.
 FT Lipid 2 N-myristoyl glycine (in host).
 SQ SEQUENCE 212 AA; 24415 MW; EC095A1B263047CF CRC64;

Query Match 43.3%; Score 78; DB 1; Length 212;
 Best Local Similarity 75.0%; Pred. No. 0.00035;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 TERGRFPVTPQVLRPMTYK 34
 : |||||
 Db 69 SEGVGFVRPQVLRPMTYK 88

RESULT 11
 ID NEF_HV1MN STANDARD; PRT; 182 AA.
 AC P05856;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
 GN NEF.
 OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 OK NCBI_TaxID=11696;
 RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Guigo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RR Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -1- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17449; AAA44858.1; -.
CC HSSP: P03406; IEFN.
CC DR HIV; M17449; NEFSNM.
CC DR InterPro; IPR001558; HIV_Nef.
CC DR Pfam; PF00469; F-Protein; 1.
CC DR ProDom; PD000031; HIV_Nef; 1.
CC KW AIDS; Myristate; GTP-binding; Lipoprotein.
CC FT Lipid 2
CC FT N-myristoyl glycine (in host) (By
CC similarity).
CC SQ SEQUENCE 182 AA; 20632 MW; 2C5B5A6E3EF13B26 CRC64;
Query Match 42.8%; Score 77; DB 1; Length 182;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 20 FVYTPQVPLRPMTYK 34
DB 70 FVYTPQVPLRPMTYK 84
RESULT 12
NEF_HV1S3 STANDARD; PRT; 205 AA.
AC P19545;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN (1)
RN SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RL J. Virol. 64:4016-4020(1990).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AY352275; AAQ17032.1; -.
CC HSSP: P03406; IEFN.
CC DR HIV; M38427; NEFSF33.

DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
DR KW AIDS; Myristate; GTP-binding; Lipoprotein.
DR FT Lipid 2
DR FT N-myristoyl glycine (in host) (By
DR similarity).
DR SQ SEQUENCE 205 AA; 23318 MW; FFB419A1C5DFC9F3 CRC64;
Query Match 42.8%; Score 77; DB 1; Length 205;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 20 FVYTPQVPLRPMTYK 34
DB 67 FVYTPQVPLRPMTYK 81
RESULT 13
NEF_HV1U4 STANDARD; PRT; 205 AA.
AC P24741;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (strain Ugandan / isolate U455)
OS (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11703;
RN (1)
RN SEQUENCE FROM N.A.
RX MEDLINE=91090981; PubMed=2265025;
RA Oram J.D., Downing R.G., Roff M., Clegg J.C.S., Serwadda D.,
RA Carswell J.W.;
RT "Nucleotide sequence of a Ugandan HIV-1 provirus reveals genetic
RT diversity from other HIV-1 isolates.";
RL AIDS Res. Hum. Retroviruses 6:1073-1078(1990).
CC -1- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC activities. It seems to down-regulate the CD4(T4) antigen.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M62320; AAA75023.1; -.
CC HSSP: P03406; IEFN.
CC DR InterPro; IPR001558; HIV_Nef.
CC DR Pfam; PF00469; F-protein; 1.
CC DR ProDom; PD000031; HIV_Nef; 1.
CC KW AIDS; Myristate; GTP-binding; Lipoprotein.
CC FT Lipid 2
CC FT N-myristoyl glycine (in host) (By
CC similarity).
CC SQ SEQUENCE 205 AA; 23253 MW; 869AB03B6E7893C4 CRC64;
Query Match 42.8%; Score 77; DB 1; Length 205;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 20 FVYTPQVPLRPMTYK 34
DB 68 FVYTPQVPLRPMTYK 82
RESULT 14
NEF_HV1EL STANDARD; PRT; 206 AA.
AC P04604;
DT 13-AUG-1987 (Rel. 05, Created)

```

DT 13-AUG-1987 (Rel. 05, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF
OS Human immunodeficiency virus type 1 (EDI isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).
RN [2]
RP POST-TRANSLATIONAL MODIFICATIONS, AND FUNCTION.
RX MEDLINE=88039140; PubMed=3118220;
RA Guy B., Kieny M.-P., Riviere Y., le Peuch C., Dett K., Girard M.,
RT "HIV F3' ORF encodes a phosphorylated GTP-binding protein resembling
an oncogene product.";
RL Nature 330:266-269(1987).
CC -I- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
activities. It seems to down-regulate the CD4(T4) antigen.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K03454; AAA44330.1; -
DR EMBL; A07108; CAA00617.1; -
DR HSSP; P03406; 1EEN.
DR HIV; K03454; NEPSFELI.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-Protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; Myristate; GTP-binding; Lipoprotein.
FT LIPID 2
FT SEQUENCE 206 AA; 23612 MW; 4CFE9F18AEB503C CRC64;

Query Match 42.8%; Score 77; DB 1; Length 206;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
Db 69 PPVTPQVPLRPMTYK 83

RESULT 15
NEF_HV1S1 STANDARD; PRT; 208 AA.
ID NEF_HV1S1
AC P19546;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Negative factor (F-protein) (27 kDa protein) (3'ORF).
GN NEF.
OS Human immunodeficiency virus type 1 (SP162 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -I- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating

```

```

CC activities. It seems to down-regulate the CD4(T4) antigen.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M65024; AAA5073.1; -
DR HSSP; P03406; 1EEN.
DR HIV; M38428; NEPSF162.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-Protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; Myristate; GTP-binding; Lipoprotein.
FT LIPID 2
FT SEQUENCE 208 AA; 23684 MW; A0B1007D14E4E32 CRC64;

Query Match 42.8%; Score 77; DB 1; Length 208;
Best Local Similarity 93.3%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 PPVTPQVPLRPMTYK 34
Db 70 PPVTPQVPLRPMTYK 84

Search completed: July 20, 2004, 06:31:02
Job time : 7.68852 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:09:43 ; Search time 34.2787 Seconds

(without alignments)
312.953 Million cell updates/secTitle: US-09-673-166b-276
Perfect score: 180
Sequence: 1 GRQYKANSKFIGITGRGRFPVTPQVPLRPMTYK 34Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.5	48.1	142	15	Q8JB99 human immun
2	86	47.8	141	15	Q8JB98 human immun
3	86	47.8	142	15	Q8JB95 human immun
4	85.5	47.5	208	15	Q74912 human immun
5	85	47.2	142	15	Q8JB97 human immun
6	85	47.2	142	15	Q8JB96 human immun
7	85	47.2	206	15	Q8JB98 human immun
8	84	46.7	208	15	Q8JB98 human immun
9	83.5	46.4	207	15	P90287 human immun
10	83	46.1	90	15	Q41645 human immun
11	83	46.1	118	15	Q41584 human immun
12	83	46.1	121	15	Q79786 human immun
13	83	46.1	123	15	Q85587 human immun
14	83	46.1	153	15	Q74916 human immun
15	83	46.1	176	15	Q74918 human immun
16	83	46.1	202	15	Q8Q617 human immun

17	83	46.1	202	15	Q9QPN3 human immun
18	83	46.1	204	15	Q9QSC4 human immun
19	83	46.1	204	15	Q9WIP2 human immun
20	83	46.1	204	15	Q9WIP3 human immun
21	83	46.1	204	15	Q74920 human immun
22	83	46.1	204	15	Q74921 human immun
23	83	46.1	204	15	Q74922 human immun
24	83	46.1	204	15	Q75628 human immun
25	83	46.1	204	15	Q75628 human immun
26	83	46.1	205	15	Q9TH66 human immun
27	83	46.1	205	15	Q9TH66 human immun
28	83	46.1	205	15	Q9TH66 human immun
29	83	46.1	206	15	Q9QXW9 human immun
30	83	46.1	206	15	Q9QXW9 human immun
31	83	46.1	206	15	Q9QXW9 human immun
32	83	46.1	206	15	Q9QXW9 human immun
33	83	46.1	206	15	Q9QXW9 human immun
34	83	46.1	206	15	Q9QXW9 human immun
35	83	46.1	206	15	Q9QXW9 human immun
36	83	46.1	206	15	Q9QXW9 human immun
37	83	46.1	206	15	Q9QXW9 human immun
38	83	46.1	206	15	Q9QXW9 human immun
39	83	46.1	206	15	Q9QXW9 human immun
40	83	46.1	206	15	Q9QXW9 human immun
41	83	46.1	206	15	Q9QXW9 human immun
42	83	46.1	206	15	Q9QXW9 human immun
43	83	46.1	206	15	Q9QXW9 human immun
44	83	46.1	206	15	Q9QXW9 human immun
45	83	46.1	206	15	Q9QXW9 human immun

ALIGNMENTS

RESULT 1

Q8JB99 PRELIMINARY; PRT; 142 AA.
ID Q8JB99
AC Q8JB99;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.,
RT "Extensive variation of CXCR4- and CCR5-dependent human
RT immunodeficiency virus type 1 in a patient with acute infection.",
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEP HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
CC EMBL; AB086306; BAC02668.1; -.
DR GO; GO:0005525; F-GTP binding; IRA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 142
SQ SEQUENCE 142 AA; 15727 MW; 152FE51521B88A13 CRC64;

Query Match 48.1%; Score 86.5; DB 15; Length 142;
Best Local Similarity 58.1%; Pred. No. 0.0001;
Matches 18; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

QY 4 YKANSKFIGITGRGRFPVTPQVPLRPMTYK 34
DB 57 WKRAQER-----EKVGFPVTPQVPLRPMTYK 82

RESULT 2

08J98 PRELIMINARY; PRT; 141 AA.

AC 08J98; 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 25, Last annotation update)

DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;

RT "Extensive variation of CXCR4- and CCR5-dependent human immunodeficiency virus type 1 in a patient with acute infection.";

RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY SIMILARITY).

CC EMBL; AB086307; BAC02669.1; -

DR GO; GO:0005525; F:GTP binding; IEA.

DR InterPro; IPR001558; HIV_Nef.

DR Pfam; PF00469; F-Protein; 1.

DR Prodom; PD000031; HIV_Nef; 1.

DR AIDS; GTP-binding; Lipoprotein; Myristate.

FT NON_TER

SC SEQUENCE 141 AA; 15544 MW; 2DD72C4B9C27097 CRC64;

Query Match 47.8%; Score 86; DB 15; Length 141;

Best Local Similarity 84.2%; Pred. No. 0.00012;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34

DB 64 EKVSFPVTPQVPLRPMYTK 82

RESULT 3

08J95 PRELIMINARY; PRT; 142 AA.

AC 08J95; 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 25, Last annotation update)

DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;

RT "Extensive variation of CXCR4- and CCR5-dependent human immunodeficiency virus type 1 in a patient with acute infection.";

RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY SIMILARITY).

CC EMBL; AB086310; BAC02672.1; -

DR GO; GO:0005525; F:GTP binding; IEA.

DR InterPro; IPR001558; HIV_Nef.

DR Pfam; PF00469; F-protein; 1.

DR Prodom; PD000031; HIV_Nef; 1.

DR AIDS; GTP-binding; Lipoprotein; Myristate.

FT NON_TER

SC SEQUENCE 142 AA; 15827 MW; 09C2333F280A0FDF CRC64;

Query Match 47.8%; Score 86; DB 15; Length 142;

Best Local Similarity 84.2%; Pred. No. 0.00012;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34

DB 64 EKVSFPVTPQVPLRPMYTK 82

RESULT 4

074912 PRELIMINARY; PRT; 208 AA.

AC 074912; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Nef (Negative factor) (F-protein) (27 kDa protein).

GN NEF.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=96400183; PubMed=8806559;

RA Rather L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,

RA Hahn B., Powderly W., Arens M.;

RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects at different stages of disease.";

RT Virology 223:245-250(1996).

CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY SIMILARITY).

CC EMBL; U04452; AA838204.1; -

DR HSSP; P03406; 1EFN.

DR GO; GO:0005525; F:GTP binding; IEA.

DR InterPro; IPR001558; HIV_Nef.

DR Pfam; PF00469; F-Protein; 1.

DR Prodom; PD000031; HIV_Nef; 1.

DR AIDS; GTP-binding; Lipoprotein; Myristate.

SC SEQUENCE 208 AA; 23596 MW; 04B09B3B594DBA0 CRC64;

Query Match 47.5%; Score 85.5; DB 15; Length 208;

Best Local Similarity 58.1%; Pred. No. 0.00022;

Matches 18; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 7 ANSKFRTGRGR--FPVTPQVPLRPMYTK 34

DB 54 ADCAWEAQEGEGVFPVTPQVPLRPMYTK 84

RESULT 5

08J97 PRELIMINARY; PRT; 142 AA.

AC 08J97; 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;

RT "Extensive variation of CXCR4- and CCR5-dependent human immunodeficiency virus type 1 in a patient with acute infection.";

RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY SIMILARITY).

CC EMBL; AB086308; BAC02670.1; -

DR GO; GO:0005525; F:GTP binding; IEA.

```
DR InterPro; IPR001558; HIV Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 142
SQ SEQUENCE 142 AA; 15715 MW; 9BC3275290AEFB7D CRC64;

Query Match 47.2%; Score 85; DB 15; Length 142;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 6
ID Q8UE96 PRELIMINARY; PRT; 142 AA.
AC Q8UE96;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Negative factor) (F-protein) (27 kDa protein) (Fragment).
DE (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi H., Suzuki Y., Hoshino H., Daar E.S., Koyanagi Y.;
RT "Extensive variation of CXCR4- and CCR5-dependent human
RT immunodeficiency virus type 1 in a patient with acute infection.";
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; AB066309; BAC02671.1; -
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 142
SQ SEQUENCE 142 AA; 15715 MW; 9BC3275290AEFB7D CRC64;

Query Match 47.2%; Score 85; DB 15; Length 142;
Best Local Similarity 84.2%; Pred. No. 0.00017;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 7
ID Q85588 PRELIMINARY; PRT; 206 AA.
AC Q85588;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-lymphotropic virus type III (HTLV III) 3' ORF HXB3 RNA (Negative
DE factor) (F-protein) (27 kDa protein).
GN NEF.
OS Aids-associated retrovirus.
OC Viruses; Retroid viruses; Retroviridae.
OX NCBI_TaxID=11966;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86067228; PubMed=2999715;
RA Ratner L., Starcich B., Josephs S.F., Hahn B.H., Reddy E.P.,
```

```
RA Livak K.J., Petteway S.R.Jr., Pearson M.L., Haseltine W.A., Arya S.K.,
RA Wong-staal F.;
RT "Polymorphism of the 3' open reading frame of the virus associated
RT with the acquired immune deficiency syndrome, human T-lymphotropic
RT virus type III ";
RL Nucleic Acids Res. 13:8219-8229(1985).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; X03188; CA26947.1; -
DR HSP; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 206 AA; 23419 MW; 3B25EB332A479A46 CRC64;

Query Match 47.2%; Score 85; DB 15; Length 206;
Best Local Similarity 84.2%; Pred. No. 0.00026;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 64 EKVGFPVTPQVPLRPMYTK 82

RESULT 8
ID Q91064 PRELIMINARY; PRT; 208 AA.
AC Q91064;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein) (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Saragosti S., Peeters M.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; AJ233029; CA13510.1; -
DR HSP; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KM AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 208
SQ SEQUENCE 208 AA; 23839 MW; F96EBA1A82BE03C2 CRC64;

Query Match 46.7%; Score 84; DB 15; Length 208;
Best Local Similarity 84.2%; Pred. No. 0.00037;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 16 ERGRFPVTPQVPLRPMYTK 34
   |:|||||
Db 66 EGVRFVTPQVPLRPMYTK 84

RESULT 9
ID P90287 PRELIMINARY; PRT; 207 AA.
AC P90287;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein).
```

```

GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAM184;
RX MEDLINE=96036482; PubMed=7483282;
RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,
RA Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;
RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes
RT from primary virus cultures using the polymerase chain reaction.";
RL Virology 213:80-86(1995).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=ZAM184;
RX MEDLINE=97213928; PubMed=9060617;
RA Salminen M.O., Carr J.K., Robertson D.L., Hegerich P., Gotte D.,
RA Koch C., Sanders-Buell E., Gao F., Sharp P.M., Hahn B.H., Burke D.S.,
RA McCutchan F.E.;
RT "Evolution and probable transmission of intersubtype recombinant human
RT immunodeficiency virus type 1 in a Zambian couple.";
RL J. Virol. 71:2647-2655(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=ZAM184;
RA Salminen M.O., Carr J.K., Robertson D.L., Hegerich P., Gotte D.,
RA Koch C., Sanders-Buell E., Gao F., Sharp P.M., Hahn B., Burke D.S.,
RA McCutchan F.E.;
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL: U86780; AAC57010.1; -.
DR HSSP: P03406; 1EFN.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 207 AA; 23358 MW; A267097BEF3A69F7 CRC64;

Query Match 46.4%; Score 83.5; DB 15; Length 207;
Best Local Similarity 63.0%; Pred. No. 0.00044;
Matches 17; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

QY 11 FVTPQVPLRPMTYK 34
Db 57 WLGAEHGEVGFPPKQVFLRPMTYK 83

RESULT 10
ID 041645 PRELIMINARY; PRT; 90 AA.
AC 041645;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein) (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C18;
RX MEDLINE=98105804; PubMed=9445059;
RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
RA Walker B.D., Neumann A.U., Vermond S.H., Westecky J., Jackson S.,
RA Penamore E., Cao Y., Gao F., Kalams S., Kunstman K.J., McDonald D.,
RA McWilliams N., Trkola A., Moore J.P., Wolinsky S.M.;
RT "Immunological and virological analyses of persons infected by human
RT immunodeficiency virus type 1 while participating in trials of

```

```

RT recombinant gp120 subunit vaccines.";
RL J. Virol. 72:1552-1576(1998).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL: U84886; AAC58933.1; -.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 90
SQ SEQUENCE 90 AA; 9794 MW; A3A713557855DF3A CRC64;

Query Match 46.1%; Score 83; DB 15; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
Db 68 FVTPQVPLRPMTYK 82

RESULT 11
ID 041584 PRELIMINARY; PRT; 118 AA.
AC 041584;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein) (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C13;
RX MEDLINE=98105804; PubMed=9445059;
RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
RA Walker B.D., Neumann A.U., Vermond S.H., Westecky J., Jackson S.,
RA Penamore E., Cao Y., Gao F., Kalams S., Kunstman K.J., McDonald D.,
RA McWilliams N., Trkola A., Moore J.P., Wolinsky S.M.;
RT "Immunological and virological analyses of persons infected by human
RT immunodeficiency virus type 1 while participating in trials of
RT recombinant gp120 subunit vaccines.";
RL J. Virol. 72:1552-1576(1998).
CC -1- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL: U84833; AAC58872.1; -.
DR GO: GO:0005525; F:GTP binding; IEA.
DR InterPro: IPR001558; HIV_Nef.
DR Pfam: PF00469; F-protein; 1.
DR ProDom: PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON TER 118
SQ SEQUENCE 118 AA; 13186 MW; DD529808140327C2 CRC64;

Query Match 46.1%; Score 83; DB 15; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FVTPQVPLRPMTYK 34
Db 70 FVTPQVPLRPMTYK 84

RESULT 12
Q79786 PRELIMINARY; PRT; 121 AA.
ID Q79786
AC Q79786;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```



```
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef protein (Negative factor) (27 kDa protein).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCRnef6;
RX MEDLINE=92385154; PubMed=1515209;
RA Harris M., Hislop S., Patelinasos P., Neil J.C.;
RT "In vivo derived HIV-1 NEF genes are heterogeneous and lack detectable
RT nucleotide binding activity."
RL AIDS Res. Hum. Retroviruses 8:537-543(1992).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; X63045; CAA44771.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 121 AA; 13506 MW; 0B9226A9052A2CC7 CRC64;

Query Match
Best Local Similarity 46.1%; Score 83; DB 15; Length 121;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMYTK 34
DB 66 FPVTPQVPLRPMYTK 80
|||||
RESULT 13
ID Q85587 PRELIMINARY; PRT; 123 AA.
AC Q85587;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-Lymphocytic virus type III (HTLV-III) 3'ORF HXB2 RNA (Negative
DE factor) (F-protein) (27 kDa protein).
GN NEF.
OS Aids-associated retrovirus.
OC Viruses; Retroid viruses; Retroviridae.
OX NCBI_TaxID=11966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067228; PubMed=2999715;
RA Ratner L., Starcich B., Josephs S.F., Hahn B.H., Reddy E.P.,
RA Livak K.J., Petteway S.R.Jr., Pearson M.L., Haseltine W.A., Arya S.K.,
RA Wong-staal F.;
RT "Polymorphism of the 3' open reading frame of the virus associated
RT with the acquired immune deficiency syndrome, human T-lymphotropic
RT virus type III."
RL Nucleic Acids Res. 13:8219-8229(1985).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; X03187; CAA26946.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Lipoprotein; Myristate.
SQ SEQUENCE 123 AA; 13593 MW; ABB6AC3BE7A2508A CRC64;

Query Match
Best Local Similarity 46.1%; Score 83; DB 15; Length 123;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 20 FPVTPQVPLRPMYTK 34
DB 68 FPVTPQVPLRPMYTK 82
|||||
RESULT 14
ID Q74916 PRELIMINARY; PRT; 153 AA.
AC Q74916;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96400183; PubMed=8806559;
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
RT at different stages of disease."
RL Virology 223:245-250(1996).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; U44459; AAB38209.1; -.
DR HSSP; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR001558; HIV_Nef.
DR Pfam; PF00469; F-protein_1.
DR ProDom; PD000031; HIV_Nef; 1.
DR AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 153
SQ SEQUENCE 153 AA; 16939 MW; BAE22A2A21305FDD CRC64;

Query Match
Best Local Similarity 46.1%; Score 83; DB 15; Length 153;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMYTK 34
DB 68 FPVTPQVPLRPMYTK 82
|||||
RESULT 15
ID Q74918 PRELIMINARY; PRT; 176 AA.
AC Q74918;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nef (Negative factor) (F-protein) (27 kDa protein)
DE (Fragment).
GN NEF.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96400183; PubMed=8806559;
RA Ratner L., Joseph T., Bandres J., Ghosh S., Heyden N.V., Templeton A.,
RA Hahn B., Powderly W., Arens M.;
RT "Sequence heterogeneity of Nef transcripts in HIV-1-infected subjects
RT at different stages of disease."
RL Virology 223:245-250(1996).
CC -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4 (T4) ANTIGEN (BY
CC SIMILARITY).
DR EMBL; U44461; AAB38211.1; -.
```

DR HSSE; P03406; IEFN.
DR GO; GO:0005525; F:GTP binding; IEA.
DR InterPro; IPR01558; HIV_Nef.
DR Pfam; PF00469; F-protein; 1.
DR ProDom; PD000031; HIV_Nef; 1.
KW AIDS; GTP-binding; Lipoprotein; Myristate.
FT NON_TER 176 176
SQ SEQUENCE 176 AA; 19286 MW; 8EE2DFCBF924430B CRC64;

Query Match 46.1%; Score 83; DB 15; Length 176;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 FPVTPQVPLRPMITYK 34
|||
Db 68 FPVTPQVPLRPMITYK 82

Search completed: July 20, 2004, 06:33:19
Job time : 35.2787 secs

OM protein - protein search, using sw model

Run on: July 20, 2004, 06:07:17 ; Search time 50.7213 Seconds

(without alignments)
189.400 Million cell updates/sec

Title: US-09-673-166B-276
Perfect score: 180

Sequence: 1 GRQYIKANSKFIGITERGRFPVTPQVPLRPMYK 34

Scoring table: BLOSUM62

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	97	53.9	28	2	AAV53552	AAV53552 Lipopept.
2	89	49.4	280	2	AAW71642	AAW71642 HIV POL/I
3	89	49.4	280	2	AAW71363	AAW71363 HIV POL/I
4	89	49.4	280	4	AAW98912	AAW98912 Synthetic
5	87.5	48.6	25	4	AAW84738	AAW84738 Amino acid
6	86	47.8	29	2	AAV53551	AAV53551 Lipopept.
7	85	47.2	206	2	AAW61515	AAW61515 Sequence
8	85	47.2	206	2	AAW89326	AAW89326 HIV-1 nef
9	85	47.2	206	2	AAW63323	AAW63323 Human lyn
10	84	46.7	505	6	ABR44739	ABR44739 Plasmid I
11	84	46.7	505	6	ABR44747	ABR44747 Plasmid I
12	84	46.7	1067	6	ABR44751	ABR44751 Plasmid I
13	84	46.7	1067	6	ABR44753	ABR44753 Plasmid I
14	83	46.1	30	4	AAW81115	AAW81115 A polypept
15	83	46.1	30	4	AAW81099	AAW81099 Peptide C
16	83	46.1	30	4	AAW81009	AAW81009 Polypept
17	83	46.1	33	2	AAV26605	AAV26605 HIV-deri
18	83	46.1	33	5	AAW83185	AAW83185 Lipopept
19	83	46.1	37	2	AAW53389	AAW53389 Universat
20	83	46.1	38	2	AAW74069	AAW74069 Supertan
21	83	46.1	38	2	AAV39888	AAV39888 HIV Negat
22	83	46.1	48	1	AAW42695	AAW42695 p14 of n
23	83	46.1	48	1	AAW92268	AAW92268 Peptide I
24	83	46.1	83	4	AAW94760	AAW94760 nef CTL1
25	83	46.1	83	4	AAW98916	AAW98916 HIV Nef-1

	XX	26	83	46.1	83	5	AAB83193	AbB3193 HIV cyto
	XX	27	83	46.1	123	6	AAB86000	AbB6000 Amino acti
	XX	28	83	46.1	142	6	ABR44742	AbR44742 Plasmid 7
	XX	29	83	46.1	151	5	AAM52395	Aam52395 Viral fus
	XX	30	83	46.1	206	2	AAR38893	Aar38893 Nef prote
	XX	31	83	46.1	206	2	AAAG90179	AAg90179 HTLV-III
	XX	32	83	46.1	206	3	AAAB10054	AAb10054 HIV-1 nef
	XX	33	83	46.1	206	3	AAAY0795	AAy0795 Human NEP
	XX	34	83	46.1	206	5	AAAE2336	AAe2336 Viral fus
	XX	35	83	46.1	206	6	ABUE6328	ABuE6328 Human Lym
	XX	36	83	46.1	215	2	AAAY02349	AAy02349 A repres
	XX	37	83	46.1	215	4	AAAG63231	AAg63231 Amino aci
	XX	38	83	46.1	215	7	ADBE7687	AdBe7687 HIV-1 nef
	XX	39	83	46.1	216	1	AAP60423	APa60423 Sequence
	XX	40	83	46.1	216	5	AAOI9380	AOI9380 Lymphaden
	XX	41	83	46.1	229	5	AAAE2337	AAe2337 Viral fus
	XX	42	83	46.1	302	2	AAAY02357	AAy02357 A repres
	XX	43	83	46.1	302	2	AAAY02351	AAy02351 A repres
	XX	44	83	46.1	302	4	AAAG63239	AAg63239 Amino aci
	XX	45	83	46.1	302	4	AAAG63233	AAg63233 Amino aci
 ALIGNMENTS								
RESULT 1								
AAV53552	ID	AAV53552 standard; protein; 28 AA.						
XX	AC	AAV53552;						
XX	DT	18-JAN-2000 (first entry)						
DE		Lipopeptide #3.						
KM		Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;						
KW		electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;						
KX		human immunodeficiency virus; hepatitis B virus; papilloma virus;						
XX		melanoma; malaria; parasite.						
OS		Synthetic.						
OS		Homo sapiens.						
XH		Key	Location/Qualifiers					
FH		Modified-site	1	/note= "contains palmitoyl residue attached at the N-				
FT				terminus and on the epsilon carbon of the side chain"				
XX		FR2776926-A1.						
XX		08-OCT-1999.						
XX		07-APR-1998;	98PR-00004323.					
XX		07-APR-1998;	98PR-00004323.					
XX		(INRM) INSERM INST NAT SANTE & RECH MEDICALE.						
PA		(CNRS) CNRS CENT NAT RECH SCI.						
PA		(INSP) INST PASTEUR LILLE.						
PI		Le Gal FA, Guillet JG, Gahery SH, Gras ME, Melnyk O, Tartar A,						
XX		WPI, 1999-583113/50.						
PT		New lipopeptide containing lipid regions and two epitopes, all separated						
XX		by peptide spacers that impart hydrophilicity, useful in vaccines.						
PS		Example 1; Page 9; 35pp; French.						

CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the lipopeptide
CC is hydrophilic. This peptide represents an example of them lipopeptide of
CC the invention. It contains 2 lipid residues attached at the N-terminal
CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
CC Ser and Ala-Ala-Ala. Peptides AAV5301-Y53549 represent peptide epitopes
CC used in the generation of the lipopeptides. These are used in therapeutic
CC or prophylactic compositions and vaccines to induce specific immune
CC responses against human immunodeficiency, hepatitis B or papilloma
CC viruses; p53 of melanoma or the malaria parasite

SQ Sequence 28 AA;

Query Match 53.9%; Score 97; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GROYKANSKFGITERGR 19
Db 1 GROYKANSKFGITERGR 19

RESULT 2
AAW71642
ID AAW71642 standard; protein; 280 AA.

AC AAW71642;

XX 17-OCT-2003 (revised)

DT 18-JAN-1999 (first entry)

XX HIV POL/NEF epitopes.

DE HIV POL/NEF epitopes.

XX Vector; gene therapy; vaccine; ALVAC; translation factor; K3L; E3L;
KM VCP1433; HIV; POL; nef; epitope.

XX Human immunodeficiency virus 1.

XX MO9840500-A1.

XX 17-SEP-1998.

PD 25-FEB-1998; 98WO-US003710.

XX 12-MAR-1997; 97US-00815809.

XX (VIRO-) VIROGENETICS CORP.

PA (UYAR-) UNIV ARIZONA STATE.

XX Tartaglia J, Jacobs BL, Goebel SJ, Cox WI, Gettig RR, Pincus SE;

PI Paoletti E;

XX WPI; 1998-520819/44.

DR N-PSDB; AAV58243.

XX Enhancing expression of nucleic acids in cells - by using modified

PT vectors which comprise the nucleic acid and also nucleic acid encoding a

PT translation factor.

XX Example 1; Fig 4A-C; 90pp; English.

CC This polypeptide comprises the POL and NEF epitopes of HIV-1. It is
CC encoded by VCP1433 (see AAV58243). In VCP1433, an HIV pol/nef 'string of
CC beads' cassette is placed under control of the vaccinia H6 promoter.

CC VMP6H6K3E3, containing a vaccinia H6/K3L expression cassette and

CC vaccinia E3L gene with endogenous promoter flanked by the ALVAC C6

CC insertion site sequences, was used in recombination with VCP1433 to

CC obtain VCP1452 (see AAV58244-45). K3L and E3L are vaccinia virus

CC translation factors. Novel vectors of the invention, such as ALVAC

CC vectors, include K3L and/or E3L and are used for enhancing expression of

CC gene products that they encode. The translation factors can effect

CC inhibition of eIF-2alpha phosphorylation or inhibition of protein kinase

CC PKR phosphorylation or otherwise sequester double stranded (ds) RNA,

CC increasing the effective concentration of ds RNA. The up-regulation of
CC foreign gene expression can have a profound effect on the induction of a
CC therapeutic or immunological response in a host administered or
CC inoculated with recombinants derived from these new vectors. (Updated on
CC 17-OCT-2003 to standardise OS field)

SQ Sequence 280 AA;

Query Match 49.4%; Score 89; DB 2; Length 280;
Best Local Similarity 77.3%; Pred. No. 7.9e-05;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 GITERGRFPYTPQVPLRPMTYK 34
Db 145 GLTTMGFPYTPQVPLRPMTYK 166

RESULT 3
AAW71363
ID AAW71363 standard; protein; 280 AA.

AC AAW71363;

XX 17-OCT-2003 (revised)

DT 15-FEB-1999 (first entry)

XX HIV POL/NEF epitopes.

DE HIV POL/NEF epitopes.

XX Vector; gene therapy; vaccine; ALVAC; K3L; E3L; translation factor;
KM VCP1433; HIV; POL; NEF; epitope.

XX Human immunodeficiency virus 1.

XX MO9840501-A1.

XX 17-SEP-1998.

PD 13-FEB-1998; 98WO-US002669.

XX 12-MAR-1997; 97US-00816155.

XX (VIRO-) VIROGENETICS CORP.

XX Tartaglia J, Cox WI, Gettig RR, Martinez H, Paoletti E;

PI Pincus SE;

XX WPI; 1998-520820/44.

DR N-PSDB; AAV60251.

XX Enhancing expression of nucleic acids in cells - by using modified

PT vectors which comprise nucleic acid and also nucleic acid encoding

PT translation factor and optionally translation factor.

XX Example 2; Fig 6; 102pp; English.

CC This polypeptide comprises the POL and NEF epitopes of HIV-1. It is
CC encoded by VCP1433 (see AAV60251). In VCP1433, an HIV pol/nef 'string of
CC beads' cassette is placed under control of the vaccinia H6 promoter.

CC VMP6H6K3E3, containing a vaccinia H6/K3L expression cassette and

CC vaccinia E3L gene with endogenous promoter flanked by the ALVAC C6

CC insertion site sequences, was used in recombination with VCP1433 to

CC obtain VCP1452 (see AAV60252-53). K3L and E3L are vaccinia virus

CC translation factors. Novel vectors are provided for enhanced expression of

CC at least 1 first nucleic acid molecule (NAM) in a cell having a

CC particular phenotype. The vector (e.g. NYVAC or ALVAC) is modified to

CC comprise the first NAM and at least 1 second NAM encoding a transcription

CC factor (TF), or a TF and a translation factor such as K3L and E3L, where

CC there is co-temporal expression of the first and second NAMS with respect

CC to the phenotype of the cell, and where expression of the second NAM

CC enhances expression of the first NAM by enhancing transcription or

CC transcription and translation. Also claimed is a method for increasing

CC expression of at least 1 first NAM by such a vector. The vectors can be

CC used for increasing expression of e.g. an epitope of interest, a

CC induce an immune response. The lipoprotein induces a B cell and/or a T
 CC cell response. The lipoprotein is used to provide intranasal or
 CC sublingual immunization

XX Sequence 25 AA;

Query Match 48.6%; Score 87.5; DB 4; Length 25;
 Best Local Similarity 80.0%; Pred. No. 9.1e-06;
 Matches 20; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 GROYI-KANSKFIGITERGRFPVTP 24
 |||||
 1 GROYIKKANSKFIGITERGRILKEP 25

RESULT 6

AA53551
 ID AAY53551 standard; protein; 29 AA.

XX AAY53551;

XX 18-JAN-2000 (first entry)

XX Lipopeptide #2.

XX Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
 KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
 KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
 KW melanoma; malaria; parasite.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1
 FT /note= "contains palmitoyl residue attached at the N-terminus and on the epsilon carbon of the side chain"

XX FR2776926-A1.

XX 08-OCT-1999.

XX 07-APR-1998; 98FR-00004323.

XX 07-APR-1998; 98FR-00004323.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX (INSP) INST PASTEUR LILLE.

XX Le Gal PA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;

XX WPI; 1999-583113/50.

XX New lipopeptide containing lipid regions and two epitopes, all separated
 PT by peptide spacers that impart hydrophilicity, useful in vaccines.

XX Example 1; Page 9; 35pp; French.

XX The invention relates to the generation of a lipopeptide comprising at
 CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
 CC epitope and at least one lipid residue with (i) the epitopes and lipid
 CC portion and (ii) the epitopes, being separated independently by peptide
 CC spacers. These spacers comprise sequences of amino acids which carry an
 CC overall electrical charge in neutral media to ensure that the lipopeptide
 CC is hydrophilic. This peptide represents an example of them lipopeptide of
 CC the invention. It contains 2 lipid residues attached at the N-terminal
 CC end of the molecule and 2 epitopes separated by the spacer residues Ser-
 CC Ser and Ala-Ala-Ala. Peptides AAY53301-Y53549 represent peptide epitopes
 CC used in the generation of the lipopeptides. These are used in therapeutic
 CC or prophylactic compositions and vaccines to induce specific immune
 CC responses against human immunodeficiency, hepatitis B or papilloma
 CC viruses; p53 of melanoma or the malaria parasite

XX Sequence 29 AA;

Query Match 47.8%; Score 86; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 OYKANSKFIGITERGR 19
 |||||
 4 OYKANSKFIGITERGR 20

RESULT 7

AA61515
 ID AAP61515 standard; protein; 206 AA.

XX AAP61515;

XX 27-AUG-2003 (revised)

XX 08-JUN-1991 (first entry)

XX Sequence of E' protein.

XX HIV; IAV; AIDS; diagnosis; vaccine.

XX Human T-cell lymphotropic virus 3B.

XX BP187041-A.

XX 09-JUL-1986.

XX 23-DEC-1985; 85EP-00309454.

XX 24-DEC-1984; 84US-00685272.

XX 04-DEC-1985; 85US-00805069.

XX (GETH) GENENTECH INC.

XX Capon DJ, Lasky LA;

XX WPI; 1986-177602/28.

XX N-PSDB; AAN60288.

XX Acquired immune deficiency syndrome polypeptide(s) - obcd. by molecular
 PT cloning etc. and used for diagnosis and in vaccines against virus
 PT disease.

XX Example; Fig 2; 125pp; English.

XX A comparison of AAN60287 with the cDNA of the HTLV-III genome revealed
 CC one particular clone, designated p7.11 which contained a DNA sequence
 CC encoding this peptide (AAP60308) sequence. This approx. 2.2 kilobase
 CC covers the precursor gag region and encodes, 5' to 3', p-12, p-15, p-24 a
 CC second p-15 protein, and approx. 300 extra base pairs 3' to the gag
 CC region (see AAN60288). [Updated on 27-AUG-2003 to correct OS field.]

XX Sequence 206 AA;

Query Match 47.2%; Score 85; DB 1; Length 206;
 Best Local Similarity 84.2%; Pred. No. 0.00022;
 Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 ERGRFPVTPVPRPMTYK 34
 |||||
 64 ERGFPVTPVPRPMTYK 82

RESULT 8

AA69326
 ID AAM89326 standard; protein; 206 AA.

XX AAM89326;

DT 17-OCT-2003 (revised)
 DT 01-JUN-1999 (first entry)
 XX
 DE HIV-1 nef protein sequence.
 XX
 KW Antigenic composition; primate; lentivirus; nef gene; vaccine; infection;
 KW AIDS; HIV-1; nef protein.
 XX
 OS Human immunodeficiency virus 1.
 XX
 PN US5851813-A.
 XX
 PD 22-DEC-1998.
 XX
 PF 27-JAN-1994; 94US-00186583.
 XX
 PR 12-JUL-1990; 90US-00551945.
 PR 09-JUL-1991; 91US-00727494.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Destrosiers RC;
 XX
 DR MPI; 1999-080408/07.
 DR N-PSDB; AAV81866.
 XX
 PT Lentivirus antigenic compositions - containing lentivirus with nef gene
 PT deletion.
 XX
 PS Disclosure; Fig 2A-R; 93pp; English.
 XX
 CC The invention relates to an antigenic composition comprising an isolated
 CC primate lentivirus whose genome contains an engineered non-reversible
 CC null mutation in the nef gene, or an infectious DNA clone in a carrier.
 CC The antigenic composition is used in vaccines against infection by the
 CC lentivirus, e.g. AIDS. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 206 AA;

Query Match 47.2%; Score 85; DB 2; Length 206;
 Best Local Similarity 84.2%; Pred. No. 0.00022;
 Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 16 ERGRPVTPOVPLRPMYK 34
 :|:|||||
 Db 64 EKVGPVTPOVPLRPMYK 82

RESULT 9
 ABU63323
 ID ABU63323 standard; protein; 206 AA.
 XX
 AC ABU63323;
 XX
 DT 24-SEP-2003 (first entry)
 XX
 DE Human lymphotrophic virus type III E' polypeptide.
 XX
 KW Vaccine; E'; AIDS; AIDS-associated retrovirus; blood testing;
 KW retrovirus dissemination inhibition; blood product testing.
 XX
 OS Human lymphotrophic virus type III.
 XX
 PN US6534285-B1.
 XX
 PD 18-MAR-2003.
 XX
 PF 12-APR-2000; 2000US-00547692.
 XX
 PR 24-DEC-1984; 84US-00685272.
 PR 04-DEC-1985; 85US-00805069.
 PR 08-MAY-1986; 86US-00861016.
 PR 02-AUG-1988; 88US-00227568.

PR 19-NOV-1992; 92US-00979391.
 PR 29-SEP-1993; 93US-00129009.
 PR 29-JUL-1994; 94US-00282857.
 PR 29-JUL-1998; 98US-00124596.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Berman PW, Capon DJ, Lasky LA;
 XX
 DR MPI; 2003-531076/50.
 DR N-PSDB; ACD27911.
 XX
 PT Composition useful as vaccine, has AIDS-associated retrovirus polypeptide
 PT having antigenic determinant that specifically binds to complementary
 PT antibody, and is free of other AIDS-associated polypeptides.
 XX
 PS Example 1; Fig 2; 49pp; English.
 XX
 CC The invention relates to a composition which comprises a predetermined
 CC polypeptide sequence of an AIDS-associated retrovirus (AR), or its
 CC fragment, where AR or its fragment comprises at least one antigenic
 CC determinant that specifically binds to complementary antibody, and is
 CC essentially free of other naturally occurring AIDS-associated polypeptide
 CC sequences or human proteins from cells for which AR is naturally
 CC infective. The composition is useful as diagnostic products and vaccines
 CC in the detection of and vaccination against viral aetiological agents of
 CC AIDS. The composition is also useful for inhibiting infection by AIDS-
 CC associated retrovirus and dissemination of retrovirus in infected
 CC individuals. The AR is useful in an assay system to identify compounds
 CC which inhibit AIDS-associated reverse transcriptase. The AR is also
 CC useful as diagnostic agents to detect AIDS in individuals donated blood
 CC and blood products and as immunogens in the production of neutralising
 CC antibodies which confer resistance to infection by AIDS-associated
 CC retrovirus. The present sequence represents the amino acid sequence of
 CC the human lymphotrophic virus type III E' polypeptide
 XX
 SQ Sequence 206 AA;

Query Match 47.2%; Score 85; DB 6; Length 206;
 Best Local Similarity 84.2%; Pred. No. 0.00022;
 Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 16 ERGRPVTPOVPLRPMYK 34
 :|:|||||
 Db 64 EKVGPVTPOVPLRPMYK 82

RESULT 10
 ABR44739
 ID ABR44739 standard; protein; 505 AA.
 XX
 AC ABR44739;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Plasmid p17/24opt/trnEfl insert protein sequence SEQ ID NO:55.
 XX
 KW Human immunodeficiency virus; HIV-1; vaccine; gene therapy;
 KW acquired immune deficiency syndrome; AIDS; HIV infection.
 XX
 OS Human immunodeficiency virus 1.
 OS Synthetic.
 XX
 PN WO2003025003-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 18-SEP-2002; 2002WO-EP010592.
 XX
 PR 20-SEP-2001; 2001WO-GB004207.
 PR 11-DEC-2001; 2001GB-00029604.
 PR 19-MAR-2002; 2002GB-00006462.
 XX

PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA,
 XX
 XX MPI: 2003-363125/34.
 DR N-PSDB; ACC69912.
 XX
 PT New nucleotide sequences encoding an HIV-1 gag protein and a second HIV
 PT antigen, or their fragments, useful in medicine, particularly as vaccines
 PT for treating acquired immune deficiency syndrome (AIDS) or HIV
 PT infections.
 PS Example 3; Fig 4; 116pp; English.
 XX
 XX The present invention describes a nucleotide sequence encoding an HIV-1
 CC gag protein or its fragment containing a gag epitope, and a second HIV
 CC antigen or its fragment encoding an epitope of the second HIV antigen,
 CC operably linked to a heterologous promoter. Also described: (1) a vector
 CC comprising the nucleotide sequence; (2) a pharmaceutical composition
 CC comprising the nucleotide sequence or the vector, and a pharmaceutical
 CC excipient, diluent, carrier or adjuvant; (3) an intradermal delivery
 CC device comprising the pharmaceutical composition; and (4) a process for
 CC the production of the nucleotide. HIV epitopes and antigens have virucide
 CC activity, and can be used in vaccines and gene therapy. The nucleotide
 CC sequence, vector or pharmaceutical compositions can be used in medicine,
 CC or for treating a patient suffering from or susceptible to a disease. The
 CC nucleotide sequence is also useful for manufacturing a medicament for the
 CC treatment of disease. The nucleic acid is particularly useful as DNA
 CC vaccines for treating e.g. acquired immune deficiency syndrome (AIDS) or
 CC HIV infections. The present sequence represents an HIV plasmid amino acid
 CC sequence, which is used in an example from the present invention
 SQ Sequence 505 AA;

Query Match 46.7%; Score 84; DB 6; Length 505;
 Best Local Similarity 62.1%; Pred. No. 0.00084;
 Matches 18; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

QY 6 KANSKFIGITGRGPPVTPQVPLRPMTYK 34
 || : |||||
 DB 359 KARVIMVG-----FVTPQVPLRPMTYK 381

RESULT 11
 ABR44747
 ID ABR44747 standard; protein; 505 AA.

AC ABR44747;
 XX

DT 28-JUL-2003 (first entry)

DE Plasmid p73i-GN2 insert protein sequence SEQ ID NO:71.

XX Human immunodeficiency virus; HIV-1; vaccine; virucide; gene therapy;
 KW acquired immune deficiency syndrome; AIDS; HIV infection.
 XX

OS Human immunodeficiency virus 1.
 OS Synthetic.

PN WO2003025003-A2.

PD 27-MAR-2003.

PF 18-SEP-2002; 2002WO-EP010592.

XX 20-SEP-2001; 2001WO-GB004207.
 PR 11-DEC-2001; 2001GB-00029604.

PR 19-MAR-2002; 2002GB-00006462.

PA (GLAX) GLAXO GROUP LTD.

PI Beaton A, Ertl PF, Gough GW, Lear A, Tite JP, Van Wely CA;
 XX

DR MPI: 2003-363125/34.
 DR N-PSDB; ACC69912.

XX
 PT New nucleotide sequences encoding an HIV-1 gag protein and a second HIV
 PT antigen, or their fragments, useful in medicine, particularly as vaccines
 PT for treating acquired immune deficiency syndrome (AIDS) or HIV
 PT infections.
 PS Example 11; Fig 12; 116pp; English.

XX
 XX The present invention describes a nucleotide sequence encoding an HIV-1
 CC gag protein or its fragment containing a gag epitope, and a second HIV
 CC antigen or its fragment encoding an epitope of the second HIV antigen,
 CC operably linked to a heterologous promoter. Also described: (1) a vector
 CC comprising the nucleotide sequence; (2) a pharmaceutical composition
 CC comprising the nucleotide sequence or the vector, and a pharmaceutical
 CC excipient, diluent, carrier or adjuvant; (3) an intradermal delivery
 CC device comprising the pharmaceutical composition; and (4) a process for
 CC the production of the nucleotide. HIV epitopes and antigens have virucide
 CC activity, and can be used in vaccines and gene therapy. The nucleotide
 CC sequence, vector or pharmaceutical compositions can be used in medicine,
 CC or for treating a patient suffering from or susceptible to a disease. The
 CC nucleotide sequence is also useful for manufacturing a medicament for the
 CC treatment of disease. The nucleic acid is particularly useful as DNA
 CC vaccines for treating e.g. acquired immune deficiency syndrome (AIDS) or
 CC HIV infections. The present sequence represents an HIV plasmid amino acid
 CC sequence, which is used in an example from the present invention
 SQ Sequence 505 AA;

Query Match 46.7%; Score 84; DB 6; Length 505;
 Best Local Similarity 62.1%; Pred. No. 0.00084;
 Matches 18; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

QY 6 KANSKFIGITGRGPPVTPQVPLRPMTYK 34
 || : |||||
 DB 359 KARVIMVG-----FVTPQVPLRPMTYK 381

RESULT 12
 ABR44751
 ID ABR44751 standard; protein; 1067 AA.

AC ABR44751;
 XX

DT 28-JUL-2003 (first entry)

DE Plasmid Trgn #6 insert protein sequence SEQ ID NO:80.

XX Human immunodeficiency virus; HIV-1; vaccine; virucide; gene therapy;
 KW acquired immune deficiency syndrome; AIDS; HIV infection.
 XX

OS Human immunodeficiency virus 1.
 OS Synthetic.

EH Key Location/Qualifiers

FT Misc-difference 11 /note= "encoded by CCG"

FT Misc-difference 231 /note= "encoded by AAG"

PN WO2003025003-A2.

PD 27-MAR-2003.

PF 18-SEP-2002; 2002WO-EP010592.

XX 20-SEP-2001; 2001WO-GB004207.
 PR 11-DEC-2001; 2001GB-00029604.

PR 19-MAR-2002; 2002GB-00006462.

PA (GLAX) GLAXO GROUP LTD.

XX


```
DR WPI; 2001-064173/08.
XX
XX New polypeptidic fragments from the p53 protein, useful for treatment or
PT prevention of cancer, e.g. of breast or colon.
PT
XX
XX Disclosure; Page 6; 26pp; French.
XX
CC The present sequence represents a polypeptidic fragment of HIV Nef
CC protein. The specification describes polypeptidic fragments of human p53,
CC which is overexpressed in many types of cancers. The p53 polypeptidic
CC fragments bind stably to human leukocyte antigen (HLA) type molecules.
CC The p53 peptides induce a specific immune response. They induce
CC cytotoxic T cells (CTL) of cells that express the p53
CC peptides associated with appropriate HLA molecules and induce secretion
CC of cytokines (particularly interleukin (IL)-2 and IL-4, and gamma-
CC interferon) by these CTL. The p53 peptides, derivatives, nucleic acids
CC encoding them and specific antibodies are used, in compositions or
CC vaccines, to treat or prevent diseases p53-related cancers, particularly
CC of breast, colon, lung or bladder
XX
SQ Sequence 30 AA;
Query Match 46.1%; Score 83; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 FPVTPQVPLRPMTYK 34
DB 1 FPVTPQVPLRPMTYK 15
RESULT 15
AAB30938
ID AAB30938 standard; peptide; 30 AA.
XX
XX AAB30938;
AC
XX
DT 02-APR-2001 (first entry)
XX
DE Peptide comprising amino acids 68-97 of HIV Nef protein.
XX
XX Polypeptidic peptide; E6 protein; HPV; CD4 epitope;
XX T helper cell; human leukocyte antigen; HLA; immune response; cytotoxic
XX cytotoxic T cell; CTL; cytokine secretion; interleukin-2; IL-2; IL-4;
XX gamma-interferon; HPV infection; cervical neoplasia; invasive cancer;
XX vulvar intraepithelial neoplasia; Nef.
XX
XX Human immunodeficiency virus.
XX
XX FR2794371-A1.
XX
XX 08-DEC-2000.
XX
XX 07-OCT-1999; 99FR-00012511.
XX
XX 03-JUN-1999; 99FR-00007012.
XX
XX (BIOV-) BIOVECTOR THERAPEUTICS SA.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Choppin J, Bourgault VI, Guillet JG, Connan F, Ferries E;
XX WPI; 2001-064175/08.
XX
XX New polypeptidic fragments from human papilloma virus E6 and E7 proteins,
XX useful for treatment or prevention of e.g. cervical neoplasia and cancer.
XX
XX Disclosure; Page 6; 27pp; French.
XX
XX The present sequence is derived from the Nef protein of Human
XX immunodeficiency virus (HIV). The peptide is an polypeptidic fragment of
XX Nef. The specification describes polypeptidic fragments from the E6 and
XX E7 proteins of human papilloma virus (HPV). The HPV peptides include CD4
```

```
CC epitopes recognised by T helper cells. They bind stably to human
CC leukocyte antigen (HLA) type molecules. The HPV peptides induce a
CC specific immune response, particularly cytotoxic T cells, caused by cytotoxic T
CC cells (CTL). They also induce secretion of cytokines (particularly
CC interleukin-2 (IL-2) and IL-4, and gamma-interferon) by CTL. The HPV
CC peptides, their derivatives, nucleic acids encoding them and specific
CC antibodies are used, in compositions or vaccines, to treat or prevent
CC diseases associated with HPV infection, e.g. cervical or vulvar
CC intraepithelial neoplasia and invasive cancer of the cervix uteri. The
CC antibodies are also useful for in vitro diagnosis of these diseases
XX
SQ Sequence 30 AA;
Query Match 46.1%; Score 83; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 FPVTPQVPLRPMTYK 34
DB 1 FPVTPQVPLRPMTYK 15
```

Search completed: July 20, 2004, 06:30:29
Job time : 51.7213 secs